Development of a western hemlock looper *Lambdina fiscellaria lugubrosa* hazard rating system for Interior British Columbia using discriminant function analysis and logistic regression.

by

Neil Borecky
Hons. BA, McMaster University, 1996

A Thesis Submitted in Partial Fulfilment of the
Requirements for the Degree of

MASTERS OF SCIENCE

In the Department of Biology

We accept this thesis as conforming
to the required standard.

[redacted]

Dr. Richard A. Ring, Academic Supervisor (Department of Biology)

[redacted]

Dr. Imre S. Otvos, Research Supervisor, (CFS, Department of Biology)

[redacted]

Dr. Barbara Hawkins, Department Member (Department of Biology)

[redacted]

Dr. Andrew Fall, External Examiner, (University of Simon Fraser)

[redacted]

Abstract

The western hemlock looper (WHL), *Lambdina fiscellaria lugubrosa* Hulst, is a major defoliator in British Columbia. The purpose of this thesis is to characterize the susceptibility of forest stands to defoliation by the investigative use of discriminant function analysis and logistic regression techniques upon forest stand and climatic variables. The ability to detect susceptibility of forests to WHL attacks is, at present, scale dependent. Predictions at an inter-watershed scale were far stronger than within watersheds, at the stand-level scale. Defoliated watersheds tend to be, on average, 1.8 ° C cooler and 2 cm precipitation/month wetter than undefoliated watersheds within the same biogeoclimatic sub-zones. Watershed-scale hazard rating using average minimum monthly temperature climate data is an accurate predictor of susceptibility 84-85% of the time. A weak stand-level signature can be identified using age, site-index and species composition with canonical discriminant function analysis (CDA). However, the overlap in CDA-score distribution is very high. This study suggests that outbreaks have some underlying small-scale controlling effects, climate being the obvious one, however at a stand-level, outbreaks tend to be more stochastic in nature, owing to the homogeneity of forest stands within defoliated areas.

Examiners: [redacted]

Dr. Richard A. Ring, Academic Supervisor (Department of Biology)

[redacted]

Dr. Imre S. Otvos, Industrial Sponsor, (CFS, Department of Biology)

[redacted]

Dr. Barbara Hawkins, Department Member (Department of Biology)

[redacted]

Dr. Andrew Fall, External Examiner, (University of Simon Fraser)

Table of Contents

List of Figures

List of Tables

List of Equations

List of Abbreviations

air_crow	air crown closure
ANOVA	analysis of variance
ANUDEM	Australian National University Digital Elevation Models
arc_air	arcsine-transformed air crown closure
arc_ced	arcsine-transformed cedar proportion
arc_hc	arcsine-transformed hemlock and cedar proportion
arc_hem	arcsine-transformed hemlock proportion
AU	adjacent undefoliated forest stand (sample unit)
Avppt	average precipitation
Avtmax	average maximum temperature
Avtmin	average minimum temperature
BC	British Columbia
BCMOF	British Columbia Ministry of Forests
BEC	Biogeoclimatic Ecosystem Classification (equivalent to BGCZ)
β	variable coefficient
BGCZ	Biogeoclimatic zone(equivalent to BEC)
Bt	*Bacillus thuringienensis*
°C	degrees Celsius
CDED	Canadian Digital Elevation Model
CDF	canonical discriminant function
CFS	Canadian Forest Service

cm	centimetres
CW	Western red cedar (*Thuja plicata*)
D	defoliation (non-loci) forest stand (sample unit)
D_1	discriminant function one
D_2	discriminant function two
DBH	diameter at breast height
dia	diameter
DSS	decision support system
DV	defoliated valley
EHL	Eastern hemlock looper
EHLDSS	Eastern hemlock looper decision support system
ESRI	Environmental Systems Research Institute
ESSF	Englemann spruce-subalpine fir biogeoclimatic zone
et al.	*et alia*
etc.	*et cetera*
FC-1	forest cover polygons, version one
FD	Douglas-fir (*Pseudotsuga menziesii*)
FEN	Forest Ecosystem Networks
FIDS	Forest Insect and Disease Survey
FIP	Forest Information Planning Files
GIS	Geographic Information Systems
ha	hectare
HW	Western hemlock (*Tsuga heterophylla*)

ICH	Interior cedar-hemlock biogeoclimatic zone
km	kilometres
L	loci of defoliation
m	metres
Mhz	mega-hertz
mw	mild-warm
N	North
°N	Geographic degrees north in latitude
N	number of samples
NAD27	North American Datum 1927
NAD83	North American Datum 1983
NDT	Natural Disturbance Type
NPV	nuclear polyhedrosis virus
NRCAN	Natural Resources Canada
NTDB	National Topographic Database
σ	variance
per comm.	personal communication
PFC	Pacific Forestry Centre
PIBS	polyhedral inclusion bodies
R	Spearman's or Pearson's rho-coefficient of correlation.
S	Spruce (*Picea glauca, Picea engelmannii, Picea x.*)
SBDSS	Spruce budworm decision support system
SBS	Sub-boreal spruce biogeoclimatic zone

SE	South-east
Site_ind	site index
SQL	system query language
Stand_ag	stand age
Stand_he	stand height
SW	south-west
TFL	tree farm license
TSA	timber supply area
μ	mean
U.S.A.	United States of America
UV	undefoliated valley
vk	very wet-cool
WHL	western hemlock looper
wk	wet-cool
Z	elevation (metres)

Acknowledgments

I would like to thank my supervisor and employer, Dr. Imre S. Otvos for his support, encouragement and patience. Without Dr. Otvos' generosity and vision this project would not have been possible. I am indebted to him for this opportunity. Additionally I would like to thank my academic advisor, Dr. Richard Ring for his advice and teaching me about bugs and the spirit of graduate school. Thank you also to my committee members, Dr. Barbara Hawkins from the University of Victoria and my external examiner, Dr. Andrew Fall from Simon Fraser University. I would like to acknowledge the support and valuable advice from Dr. David Gray of the Atlantic Forestry Centre . Extensive support has been provided by the Pacific Forestry Centre in Victoria, and the staff there including; Dennis Clarke, Dr. Steen Magnusen, Dr. Hugh Barclay, Dr. Allan Carroll, Nicholas Conder, and of course Andrea, Hollywood, Liz Cameron, Tim Boulton and Maarten Voordouw. My family deserves my thanks for their support and use of the cabin in the woods to writing this thesis.

Additionally I would like to thank my sponsors for their financial support:

Natural Resources Canada-Canadian Forestry Service

Science Council of British Columbia, GREAT Awards committee

Dr. Lorraine Maclauchlan, B.C. Ministry of Forests, Kamloops Forest Region

Dr. Art Stock, B.C. Ministry of Forests, Nelson Forest Region

Barry Wagner, Downie St. Sawmills Ltd.

Robert Johnson, Riverside Forest Products Ltd.

Cameron Leitch , Pope and Talbot Limited

Del Williams, Revelstoke Community Forest

1 Introduction

1.1 Statement of problem

The western hemlock looper, *Lambdina fiscellaria lugubrosa* Hulst, (Lepidoptera: Geometridae), is a major cause of defoliation damage in British Columbia. The purpose of this thesis is to characterize the susceptibility of forest stands to defoliation by the use of discriminant function analysis and logistic regression techniques upon forest stand and climatic variables. This involves developing a set of functions based upon a combination of biophysical site variables (forest-cover, elevation and climate) that best discriminate among the following groups of forest stands: adjacent undefoliated (AU), loci of defoliation (L), subsequently defoliated (D) or within a valley that is not adjacent undefoliated (UV). The functions are then applied to the characteristics of unclassified sites to predict their group membership.

Based upon the multivariate nature of the data and the many-to-many relationships between the independent and dependent variables, discriminant analysis is the most appropriate type of statistic to perform. Logistic multiple regression will be presented as a model for confirmation and comparison, as well as to provide a simpler model for probability purposes.

1.2 Hazard rating and susceptibility models

The aim of this study is to incorporate resulting models into a hazard-rating system for western hemlock looper (WHL). Typically, a hazard-rating system for insects will contain information about both the timing of an outbreak, as well as the location. The dynamic temporal component is generally known as the 'risk'. The spatial component of the hazard is referred to as the susceptibility. In this study, susceptibility and hazard may be used interchangeably.

As with most forest insect foliovores, outbreaks of WHL are somewhat difficult to predict. The WHL has an average outbreak periodicity in British Columbia of about 8 years, however the specific locations are variable (Parfett *et al.* 1995). Although the primary goal of this study is to measure the susceptibility, the risk in Interior British Columbia is currently being monitored by a network of 23 sentinel pheromone traps. The trap catches have not yet monitored the WHL populations throughout a complete outbreak cycle. Monitoring is currently continuing. There have been many decision support systems created for forest pest management, often with mixed success. The Canadian Forest Service has created a number of hazard rating and decision support systems (DSS). The Eastern Hemlock Looper Decision Support System (EHLDSS) successfully predicted the locations of eastern hemlock looper (*Lambdina fiscellaria fiscellaria*) outbreaks. However, it tended to underestimate the extent of the outbreak (Carroll, 1996). This particular model relied heavily upon history of stand defoliation and proximity to these defoliated stands for its predictions. In addition to other factors such as changing forest dynamics and stand structure (*Ibid.* 1996), the lack of quantitative analysis of pre-outbreak looper populations may have been a large factor in the underestimation of outbreak. Well recognized and quite successful is the susceptibility and risk rating system for the mountain pine beetle (*Dendroctonus ponderosae*) . The hazard-rating component measures the stand age, host basal area, stand density and elevation to measure the stand's risk of attack by beetles. Several models are deployed within this DSS (Shore and Safranyik, 1992; Shore *et al.*1996).

Gray *et al.* (1998) helped create a forecasting system using stand conditions to predict the occurrence of defoliation by spruce budworm (*Choristoneura fumiferana* Clem.). This was also combined with a number of other models, such as growth/yield - models, spraying benefits etc. to create the Spruce Budworm DSS in New Brunswick, (MacLean *et al.*, 2000).

Perhaps the most intensively-studied forest defoliator is the European gypsy moth (*Lymantria dispar*. L.) and some of the best lessons in the techniques of rating hazard and risk of defoliation come from the studies of this pest. A great deal of effort and expense has been put forth in the United States towards forecasting of gypsy moth outbreaks. It was in many of these studies that geospatial statistical techniques were introduced as a way of predicting the locations and dynamics of outbreaks. Studies have achieved some degree of success by interpolating population levels from egg-mass surveys prior to outbreaks (Gribko *et al.*, 1995; Weseloh, 1996).

Three-dimensional kriging was one of the most logical and initially promising choices of techniques. Simply put, the process of kriging involved the generation of a surface based upon point features. In three dimensions, the X-Y components of kriging, are combined with time to make a predictive assessment of future outbreak locations and timing. Hohn *et al.* (1993) used this technique based upon measures of past defoliation and egg-mass surveys in the state of Massachusetts. While the post-hoc predictions of the locations of outbreaks were fairly accurate, the actual timing of outbreaks lagged behind by a year. In addition, the 2 X 2 km resolution of this study limits its usefulness from an operational forestry perspective.

In Michigan, Gage *et al.* (1990) utilized a statewide grid of 2000 pheromone traps. This was combined with a regression model of predicted trap catches and an estimation of expected defoliation based upon the number of male moths caught in the traps. It allowed for the analysis of trends in expanding populations. Such a system would be difficult and costly to employ where access is both limited and the area to be monitored is very large, particularly in the absence of the former Forest Insect and Disease Survey (FIDS), which was disbanded in 1995.

A great deal of work was accomplished using ordinary kriging to interpolate egg-mass densities across a landscape based upon a network of intensive sampling, and infer outbreak locations (Liebhold *et al.,* 1991). These models often failed to perform reliably (Liebhold *et al.*, 1994). Liebhold and McManus (1991) concluded that *L. dispar* does not spread through the immigration of larvae from the loci of attack, but its apparently spread is actually an artifact of susceptible forest types. It may be that a similar mechanism exists in WHL as is shown through the simultaneous appearance of widespread defoliation in different loci. Susceptibility models have been applied to forest inventory to map forest susceptibility at the landscape level (Gasner *et al.*, 1993). Liebhold *et al.* (1994, 1997) calculated the susceptibility of forests to gypsy moth defoliation using correlation techniques. These susceptible forests included over 20 species of trees identified by Montgomery (1991), and were based upon field and laboratory feeding trials as well as taxonomic affinity. Liebhold *et al.* (1997) calculated the susceptibility of those forests stands on the landscape in classes of 20, 50 and 80% of the tree basal area. Landscape characteristics, particularly elevation, were found to be correlated with counts of gypsy moth in pheromone traps (Sharov *et al.*, 1997). The spread of this particular polyphagous forest pest was found to be more closely correlated with forest susceptibility than with climate (Sharov *et al.* 1999).

Gribko *et al.* (1995) used a combination of a kriged egg-mass density surface and a logistic regression model to define the probability of an outbreak occurring in the following year at a given location. This was also performed upon a 2 X 2 km resolution. It was recommended that this model be used to locate areas where finer-scale monitoring would be conducted to estimate defoliation in individual stands. It was noted that this method was an improvement upon the 3-D kriging model.

In a sense, both logistic regression and discriminant analysis are suitable candidates for determining a landscape's risk of defoliation and answer the same questions regarding forest susceptibility. Southern pine beetle (*Dendroctonus frontalis* Zimm.) outbreaks were described using logistic regression, while accounting for spatial and temporal autocorrelation (Gumpertz *et al.*, 2000). Varnier *et al.* (1998) developed a climatic-determined risk model for schleroderris disease for pines in Ontario using Dan McKenny's climatic surface model and logistic regression. The final model was parameterized around mean temperature and precipitation from the coldest quarter of the year.

Using these same climatic data, Roland *et al.* (1998) noted that long term minimum temperature for the coldest month, and the predicted growing degree days for the first six weeks of the growing season are negatively associated with outbreak duration in forest tent caterpillar (*Malacosoma disstria* Hbn.). However, forest heterogeneity accounts for more variation than either climatic variable.

There are many applications outside the realm of forest pest hazard rating that model the spatial distribution of biological phenomena. Although too numerous to list them all within the scope of this thesis, an excellent example can be drawn from a study of the distribution of red-crown crane (*Grus japonensis*) in China. While taking into account autocorrelation, it used logistic regression of principal components to simulate and predict the distribution and spread of this re-introduced species, based upon presence/absence data of the species (Wang and Tang, 1997). In many ways, this is analogous to modeling the potential distribution of outbreaks based upon the presence/absence of defoliation.

1.3 History of *Lambdina fiscellaria lugubrosa* in British Columbia

The earliest recorded signs of western hemlock looper (WHL) defoliation were recorded in Vancouver's Stanley Park from 1911 to 1914 and since that time there have been 14 outbreaks of WHL in British Columbia (BC) (Table 1) (Harris et al. 1982; Erickson, 1992; Parfett *et al.,* 1995).

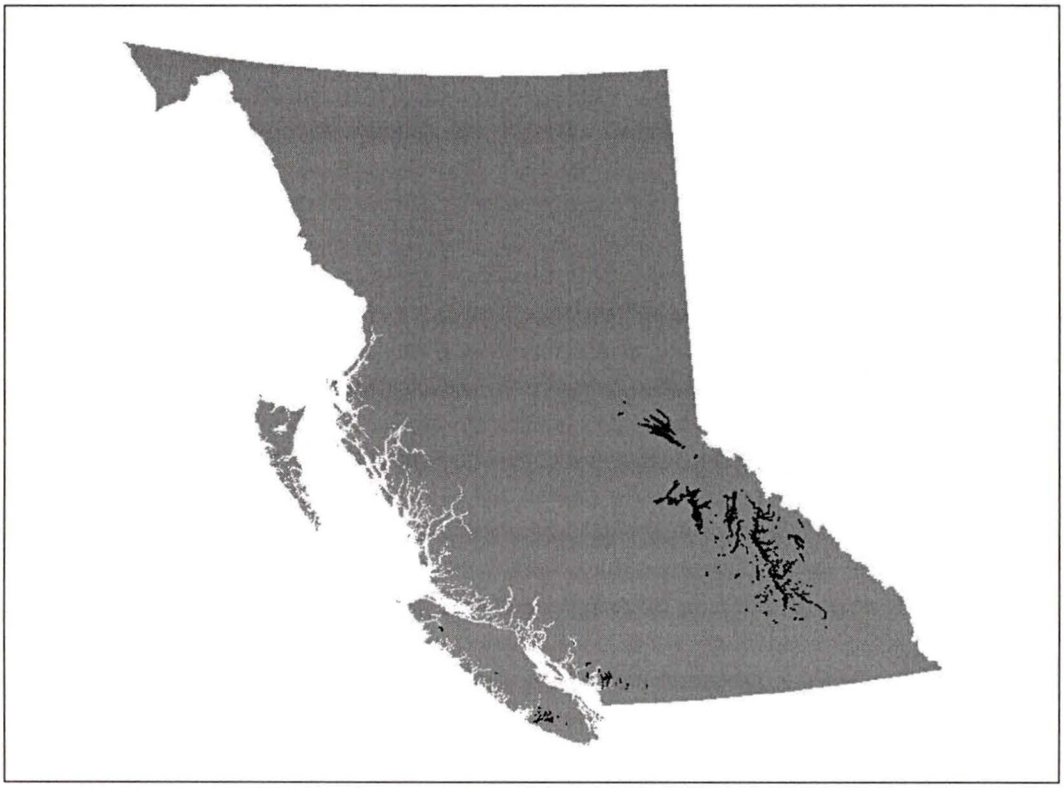


Figure 1. Area defoliated by western hemlock looper in British Columbia, 1911-1995.

Figure (1) shows the extent of outbreaks in British Columbia. A lengthy history of observation and systematic monitoring starting in the 1940s suggests that these outbreaks have increased in size, distribution and intensity. However, the use of more extensive surveillance techniques, particularly the commencement of annual aerial survey practices in 1960, may account for some of the apparent increase. In the last outbreak alone (1990-1995), 272 484 ha of trees were defoliated to some extent by WHL, with over 63 000 ha of trees killed (Borecky and Otvos, 2001). Federal monitoring of insects and pathogens affecting forest health has occurred since 1912 at some level (Van Sickle *et al.* 2001). At present, defoliation data, including WHL, is collected by the British Columbia Ministry of Forests through the use of private contractors. Monitoring of larval and adult populations is handled federally.

Reviews of past WHL activity have shown that outbreaks have occurred in the wet belts of both coastal (SW) and interior (SE) British Columbia south of 56°N latitude (Jardine, 1969; Turnquist, 1991; Parfett *et al.*, 1995). Although, not observed in BC above this latitude, WHL outbreaks have been reported in the Bradford/Bradfield River area of Alaska, but not on a large scale, indicating a range that extends considerably farther north, (Crosby and Baker,1966). Smaller outbreaks have been recorded the Coastal Western Hemlock Biogeoclimatic Zone in BC, on Vancouver Island, and within the lower Fraser Valley. Earlier outbreaks may have been more extensive than reported. On Vancouver Island, it was reported that between 1913 and 1914, large areas of hemlock were defoliated and killed by WHL (Allen and Koot, 1974). However, no mapped record of this activity exists. Anecdotal evidence or written description are the prime sources of information for these earlier outbreaks, particularly prior to 1949 (Dennis Clarke-CFS, personal communication; Parfett *et al.* 1995).

Recently, WHL defoliation has started and is occurring within the Vancouver Region watershed, in stands as young as 60 years (Kimoto and Carlson, 2001). Because the timing of attacks and extent of WHL damage is considerably different between the coastal and southern-interior areas of BC, this thesis will primarily concentrate upon the more recent and severely impacted interior region of the Province.

Outbeak Periods	Location	Size (ha)
1911-1914	SW BC	> 100
1925-1927	SW BC	1 423
1927-1929	SW BC	4 057
1937-1938	SE BC	36 532
1944-1947	SW BC	13 969
1945-1947	SE BC	89 025
1954-1955	SE BC	45 516
1958-1959	SW BC	177
1961-1964	SE BC	2 889
1969-1973	SW BC	300
1972-1976	SE BC	78 491
1982-1985	SE BC	45 438
1987	SW BC	89
1990-1995	SE BC	272 484

Table 1. Outbreak history of western hemlock looper in British Columbia (1911-1995)

The WHL has been responsible for large areas of severe defoliation causing tree mortality in recent years, notably in the federally-designated forest regions of Nelson, Cariboo, and Prince George. This spans an area encompassing the eastern and central regions of British Columbia from McBride, south to Nakusp, and east towards Kamloops. This area in BC is the focus of our study.

1.4 Hosts of western hemlock looper

WHL outbreaks characteristically begin in old-growth western hemlock (*Tsuga heterophylla* (Raf.) Sarg.), and western red cedar (*Thuja plicata* Donn), within the Interior Cedar-Hemlock (ICH) Biogeoclimatic Zone. Roughly 80% of the observed defoliated area from 1911-1995 in British Columbia has been within the Interior cedar hemlock (ICH) very wet cool (vk), ICH wet cool (wk), and ICH moist warm (mw) subzones (Borecky & Otvos, 2001; Parfett *et al.,* 1995). During outbreaks, defoliation can spread into stands that are outside these subzones, and adjacent to previously defoliated cedar-hemlock stands, (Humphreys, 1994). Defoliation has been noted in other biogeoclimatic zones, and in tree species such as Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco), true fir (*Abies spp.*), western white pine (*Pinus monticola* Dougl.), Sitka spruce (*Picea sitchensis* (Bong.) Carr.), Englemann spruce (*Picea engelmannii* Parry), white spruce (*Picea glauca* Moench), and western larch (*Larix occidentalis* Nutt.) among others, including occasional deciduous trees and shrubs (Cottrell and Monts., 1976; Harris *et al.,* 1982; Parfett *et al.*, 1995).

Overall, outbreaks appear to follow the host species distribution and consequently are limited to lower valleys and lower slopes upon both sides. Mature timber with a high proportion of hemlock appears to be the most susceptible (Wyatt, 1946, Borecky and Otvos, 2001).

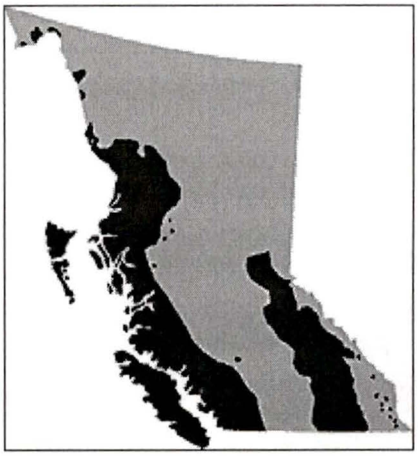


Figure 2. Range of *Tsuga heterophylla* in British Columbia (BCMOF, 2001)

Tsuga heterophylla itself is the primary host for WHL, and is an important commercial species in both the Pacific Northwest and the northern Rocky Mountain regions of North America. Its range extends from central California along the British Columbia west coast and associated islands to the Kenai Peninsula of Alaska, as well as on the western portions of the U.S. Cascade range. In the interior, it grows on the west side of the Continental Divide, from Montana in the United States, to north of Prince George, British Columbia (Hosie, 1969).

Figure 2 shows the range of *T. heterophyllya* in British Columbia.Western hemlock grows best in wetter climates, with frequent precipitation throughout the growing season. In drier regions, hemlock is confined to wet sites along valley floors, seepage areas and some north-facing slopes (Packee, 1976). In British Columbia, hemlock ranges from sea-level to about 1400 metres in elevation (Borecky, personal observation).

Western hemlock foliage is characterized by short needles, about 10-20 mm in length. The branches exhibit strong apical control, but with weak apical dominance, giving the branches a very broad appearance. The mature trees have irregular crowns with many internodal branches. Hemlock also exhibits free growth, and continues to produce sylleptic shoots throughout the growing season (Taylor, 1993).

Old-growth hemlocks are often 100-275 cm in diameter at breast height; attaining heights of 50-60 metres but occasionally up to 80 metres. Trees typically reach a maximum of 400 years in age depending on conditions, but usually do not gain in height after 300 years (Packee, 1976).

Good quality western hemlock is valued by the forest industry for construction lumber, pilings, poles and railway ties. It has good to excellent pulping characteristics and is used as a fibre source for the manufacture of paper products in thermo-mechanical, kraft and sulphite pulps (Packee, 1976; Taylor, 1993).

1.5 Western hemlock looper population monitoring

From as early as 1912 until its dissolution in 1995, the Canadian Forestry Service operated nation-wide the Forest Insect and Disease Survey (FIDS), although not always by that official name designation in early years (Van Sickle *et al.,* 2001). In cooperation with the British Columbia Ministry of Forests (BCMOF), FIDS conducted extensive province-wide surveys on the health of the forests in B.C. in each of the six former federally-designated forest regions: Vancouver, Prince Rupert, Cariboo, Kamloops, Prince George and Nelson.

There have been two main methods of monitoring WHL populations: ground sampling and aerial survey. Ground sampling, via egg, larval, pupal and adult moth censuses examined insect populations including WHL, throughout the Province. This type of census provided information in times of low, endemic population levels, as well as during defoliation events.

There have been various attempts over the years to estimate defoliation damage from larval numbers. It was estimated in British Columbia that when 40% of all larval samples collected were positive for WHL, a local outbreak may develop, and that Interior outbreaks were more severe than coastal outbreaks (Silver, 1962). Harris *et al.* (1982) reported that there was a difference between larval levels and notable damage, when comparing British Columbia's coastal and interior outbreaks. On the coast, the average number of WHL per sample could be as low as 2 or 3, and when between 31% and 34% of the samples in the area were positive, damage would be visible in the following year. The height of an outbreak on the coast would see at least 39% of samples positive for WHL. The Interior was less sensitive to larval levels, with at least 64% of samples containing an average of 8 to 10 looper per positive sample, indicated a probable defoliation event in the following year. During an outbreak year, 81% of sample plots generally would be positive for WHL larvae in a standard 3-tree beating. Turnquist (1991) noted a similar difference between coastal and interior areas; however, he estimated that only 45% of 3-tree beating samples needed to be positive for looper in the Interior in the year before an outbreak.

Logically, egg-surveys should give some insight into the number of larvae that will emerge to feed upon foliage. For example, Thompson (1958) estimated that when egg numbers reached 10 eggs/ 0.5 ft^2 of bark at mid-crown, noticeable defoliation may occur the following year. This method of egg-sampling is somewhat more labour intensive and costly than other sampling methods as it requires the tree to either be felled or scaled. The FIDS method of egg survey was based upon a count of the average number of healthy eggs per 100 grams of 'Old Man's Beard' lichen (*Usnea barbata* L.) (Turquist, 1991). The projection of defoliating population levels based upon egg surveys has been utilized with some degree of success to forecast gypsy moth (*Lymantria dispar*. L.) outbreaks (Gribko *et al.*, 1995; Weseloh, 1996). Studies of eastern hemlock looper (EHL) where the relationship between the number of eggs, the number of larvae, and then the resulting defoliation was analyzed by regression analysis. This produced significant equations but with very low R^2 values (Trial *et al.*, 1991). Even frass production has been attempted to estimate population levels (Thompson, 1958).

Light traps have been used to census adult populations. In the eastern United States, black-light (Ultra-violet) traps capturing over 500 EHL moths per night were noted to be a good estimator of populations high enough to cause visible damage to foliage in the following year (Simons *et al.*, no date). It is possible to combine light traps with a pheromone in order to fine-tune catch numbers (Jobin and Coolombe, 1992).

Pheromone trapping of adult WHL is the most effective method for broad-area population estimates. Consider that 3-tree beatings for larvae, or egg-collection results are fairly site-specific, and insect density may vary considerably from tree to tree, or even from base to crown and branch to branch. Pheromone disperses over a wide area and although it is dependent upon wind direction, dispersal pattern, dosage and other factors, the broadcast-nature of this method of sampling enables more to be said about a particular locale rather than a particular tree. It has been shown that trap catches are spatially autocorrelated at a distance of 4km, and in some cases, up to 10 km (Liang *et al.*, 1997). There are several reasons that the use of pheromone trapping has become the chosen method for the monitoring of WHL populations: 1) it allows the sampling to be less site-specific, 2) it is less labour intensive, requiring about one week to deploy and one week to collect the traps at the 23 permanent monitoring sites across the Province, and 3) pheromone trapping requires less expertise and the sampling technique experiences less variation than tree beatings or lichen collection for egg sampling (I.Otvos and N. Conder-CFS, personal communication ; Evenden *et al.*, 1995).

Placing pheromone traps for *L. fiscellaria* entails baiting a rubber septum lure with a dose between 10μg of isomeric 5, 11 dimethylheptadecane and 2,5 dimethylheptadecane mixed at a 1:1 ratio. (Evenden *et al.*, 1995). The 5,11 dimethylheptadecane is the most attractive component; however, 2,5 dimethlyheptadecane acts as a synergist. In addition it has been found that 7 dimethlyheptadecane in combination with the other two components increases the catch by seven times that of traps baited with only 5,11 dimethlyheptadecane (Krannitz, 1992). The number of male moths caught is a direct function of dose, so that exact care must be taken when administering the dosage if trap catches are to be compared from year to year. Varying trap type and pheromone dosage will produce markedly different trap catches, (Van Sickle, 1994; Carlson, *per. comm.* 2001). One of the shortcomings of this method of sampling is that the number of male moths caught in the fall of one year has not been found to be an accurate predictor of the number of larvae (or pupae) in the following generation; a key factor to consider as the larvae are the destructive agent. However, the number of males caught in one generation has been linked to the number of eggs in the subsequent generation (laid the same year of moth flight) at $R^2 = 0.60$ (Evenden *et al.*, 1995). For this reason, any forecasting program based upon pheromone trap catch numbers, will most likely be a broad-based warning of impending defoliation when moth catch numbers rise to within a certain range.

Between 1992 and 2002, a network of 23 sentinel traps have monitored looper populations around the Province. While this network has yet to follow WHL populations through a complete cycle from endemic levels to outbreak proportions and back down to endemic levels, it is anticipated that once data become available from a complete cycle, this system could be used to forecast impending outbreaks (Liang, 1997; Borecky and Otvos, 2001). There are currently strong indications of a population in ascent, however, without signs of visible defoliation or outbreak, trap catch numbers for the years prior to damage have not yet been determined in order to construct even a rough guide or warning system (Figure 3). Once adult populations have been followed through a complete cycle, it will be possible to determine a threshold level that indicates defoliation in the following year.

One interesting observation from the more recent monitoring efforts, has been that although trap counts have been up province-wide, the areas that experienced the greatest mortality in hemlock-cedar stands in the 1990-1995 outbreak of WHL, currently show some of the lowest catch numbers, particularly in the McBride area (Figure 3). This could be an indication that regions that have been denuded of host tree stands are no longer able to sustain population levels at a capacity that could eventually cause damage. This hypothesis can only be tested as the WHL population completes a full cycle. However, it underscores the importance of incorporating forest inventory data into any sort of hazard rating, as host species would intuitively be a key factor to consider.

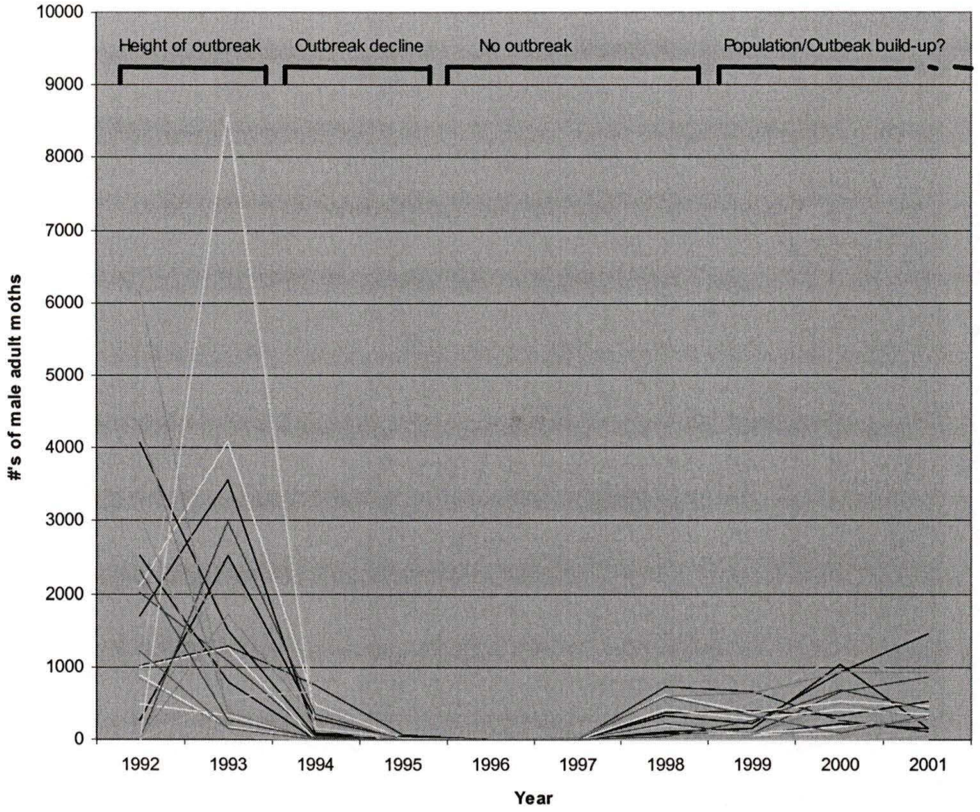


Figure 3. Pheromone trap counts for western hemlock looper at 23 locations in British Columbia, 1992-2001.

In addition to ground sampling for population levels, damage surveys from ground and later by air have provided valuable historical information about defoliation events. The FIDS-produced database includes defoliation records by WHL from 1911-1995, and contains over 165 digital maps (Parfett *et al.,* 1995). Much of the early data is based upon written descriptions of the location of the outbreaks, for example the 1911-1913 outbreak in Stanley Park. FIDS rangers' annual reports formed the basis of the digital compilation. Prior to 1960, in the interior of the Province, ground surveys of defoliation were conducted via automobile, foot, boat, and even on horseback, as well as by the occasional aerial over-flight. It is entirely probable that these early surveys were not exhaustive in their coverage of the actual outbreak extent, particularly as road networks were not as extensive as they are today.

Trees defoliated by WHL take on a characteristic red-brown hue as the needles of defoliated trees are only partially consumed in many cases. It can be distinguished from the previous year's defoliation where affected areas turn grey and the trees appear threadbare. Affected areas are sampled from the ground to confirm the identity of the pest causing the damage. From the hardcopy 'sketch maps', the data are transferred by digitization to incorporate each region, and each year into the FIDS database. Although the total WHL dataset is quite large , the period spanning 1960-1995 is both the most complete and the most precise data, as aerial surveys can logically capture more area, providing greater coverage. It may be partially due to this fact that outbreaks seem more extensive in this time period. For this reason, the 1960-1995 dataset is what we used initially, although for reasons related to the completeness of the BCMOF forest inventory, we eventually used only the 1990-1995 outbreak in our final analysis.

1.6 Life history

It is essential to understand the biology and ecology of the WHL, including host tree and pest interaction, before embarking upon any form of hazard rating. From this, clues may be extracted that explain the preference of this pest for one particular area over another.

The hemlock loopers are a complex of three sub-species, The western hemlock looper (*Lambdina fiscellaria lugubrosa* (Hulst), is a subspecies of *Lambdina fiscellaria* (Guenee). *L. fiscellaria* is a widely distributed native species of North America. The eastern sub-species *L. fiscellaria fiscellaria* (Guen.), has a range extending from Newfoundland, through the Atlantic Provinces, New England, Quebec, Ontario and some of the Lake States. In the west, the sub-species of interest, *L. fiscellaria lugubrosa*, occurs in Alaska, British Columbia, Washington and Oregon, particularly in damper regions with higher rainfall. A third sub-species, *L. somniaria* (Hulst), the western oak looper, occurs primarily on deciduous foliage in western British Columbia and Washington state and is considered to be a host-induced variant. (Raske *et al.,* 1986). All three are considered one species due to the lack of differences in genitalic and other morphological features such as setal pattern of larvae and cremaster structure of pupae (McGuffin, 1987).

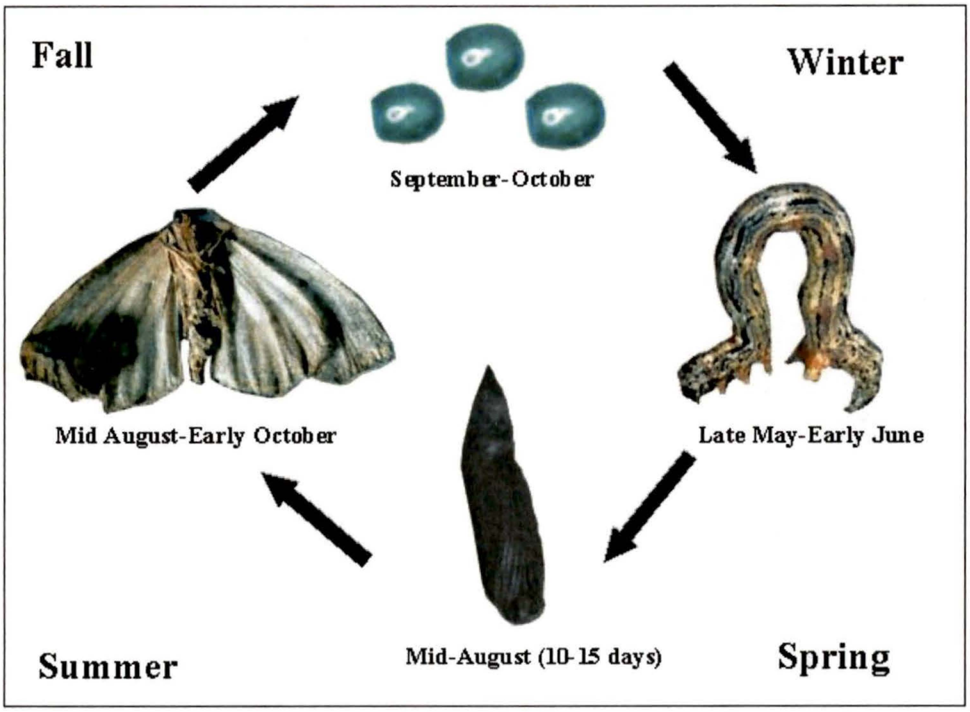


Figure 4. Life cycle of the western hemlock looper (life stages are not drawn to scale).

1.6.1 Larvae

After overwintering in the egg stage, the first larval instars begin emerging in May and early June. The larvae are 5mm to 33mm long depending upon the instar. Early instars have a black head with a longitudinally light grey striped body. Upon emergence, they begin feeding upon newer foliage in the upper crown. The mandibles of young larvae may not be developed enough to deal with tougher old foliage (Ouellet, 1983; Carroll, 1999). Later instars are mottled grey-brown with two pairs of dark spots visible on the dorsal side of each abdominal segment (Jardine, 1969). The larvae are wasteful feeders, often consuming only portions of a needle before moving onto the next. Later instars consume both new and old foliage. Feeding continues through six larval instar stages until mid to late August when the mature larvae pupate. Trees are partially or completely defoliated in a single season.

1.6.2 Pupae

Lasting from 10-15 days, the pupal stage of the *L. fiscellaria lugubrosa* occurs within the bark crevices on the trunk, on foliage, upon lichens and mosses. The pupal case is between 11-15 mm in length and is mottled greenish-brown in appearance (Jardine, 1969; Ostaff *et al.,* 1974.)

1.6.3 Adult

Adult emergence occurs in mid- to late August and the moths fly until October. Males emerge before females. Upon emergence from the pupal casing, the adults immediately fly in search of a mate. The adults are light brown to beige in colour with the characteristic geometrid triangular forewing and a smaller, more rounded rear wing. The wing-span is about 32 mm for the fore-wings. These fore-wings exhibit two short-axis wavy lines that run parallel to the body. The rear wings have only one wavy line (Jardine, 1969). The moths fly from mid to late August right into October. The greatest mating success is between two to four-day old males and with females up to four days old at a 6:1 male-female sex-ratio (Ostaff *et al.*, 1974). There appears to be a mating threshold temperature of 20 to 21°C. Mating tends not to occur mid-day under conditions of continuous light, but around dusk or in shaded or partially shaded conditions. *L. fiscellaria fiscellaria* appears to have a peak flight period one half hour after sunset, in evening temperatures above 4°C, with a second peak flight period around midnight. The majority of WHL flight ceases at dawn. Flight activity occurs in stands with a live canopy at least 3 m above the ground. Observations of between-stand dispersal occur very rarely (Raske and Farrell, 1986). Krannitz (1992) observed daily moth flight to begin one hour after sunset, with peak flight periods between two and seven hours after sunset depending upon temperature and weather conditions. Heavy rains, high winds, and adverse weather conditions tend to disrupt moth flights and mating activities. Lower temperatures increase the length of the adult stadium and the number of eggs laid. Temperatures below 12.5 °C appear to reduce oviposition (Thompson, 1956; Jardine, 1969; Erickson, 1984; Koot, 1994). These factors ought to be considered when examining climatic and weather variables for inclusion into the first stage of our hazard rating model that distinguishes valleys at risk from valleys not at risk. Conclusions drawn from our results should incorporate some of the life history mentioned above to draw a biological connection to the findings.

1.6.4 Eggs

In *L. fiscellaria lugubrosa,* oviposition occurs between September and early October, with eggs being laid primarily on strands of lichens such as 'Old Man's Beard' (*Usnea barbata*), 'Black Moss' (*Alectoria spp.*), and beneath flat lichens (*Pamelia spp.*) on the branches and trunks of trees (Campbell, 1946). The number of eggs laid are positively correlated with the pupal weight of the female (Thompson, 1956). The number of eggs laid per female has been observed to be between 1 and 177, with an average of 58 (Hopping *et al.*, 1938). The eggs are about 1mm x 0.75mm in size, smooth in texture, ovoid in shape and pale gray-green to brown in colour (Figure 4). In *L. fiscellaria fiscellaria*, Jobin and Desaulniers (1981) observed that 98% of eggs were laid on host trees, primarily within or beneath lichens, and the remaining 2% laid on old logs, stumps etc. Both host tree and lichen play an important role in ovipostion. On Anticosti Island, it was observed that the lichen preferred by looper were positively associated with stand age and density (Laflamme *et al.*, 1979). The looper eggs overwinter and larvae emerge to commence feeding in the spring. For *L. fiscellaria fiscellaria species*, larval emergence and development have been observed to be closely related to the accumulation of degree days above 3° C (Hartling and Carter, 1993).

1.7 Natural predators, parasitoids, and pathogens of *Lambdina fiscellaria lugubrosa.*

When exploring bio-physical variables that may explain the presence or absence of the WHL population fluctuations, or outbreak-levels of WHL, the life history of agents controlling looper numbers are of interest. Predators, parasitoids and pathogens have long been implicated in the decline of looper populations from high levels (Jardine, 1969; Van Sickle, 1995). There is little information on the mechanisms that limit the WHL's population . It is entirely possible that a fine balance exists between parasitoids/pathogens, and that looper fecundity keeps the populations at endemic levels. Similar or related to the concept of a 'climatic release', negative changes in the abundance or success of looper-targeted attacks within parasite/pathogens may allow temporary relief from this natural process of population suppression. This would allow the population to increase exponentially over a few generations and lead to outbreak conditions; a situation that would remain until the natural control factors catch up and once again suppress the pest population. While the individual life histories of most WHL pathogens and parasitoids are not described in detail, the species themselves are presented as agents of population control for the WHL. A large body of this research has been conducted on the eastern hemlock looper. Since the two species are so similar, it seems reasonable to draw parallels with similar control agents, particularly if the agents are closely related or the same in the two locations. It is expected that the results of our discriminant analysis technique may present opportunity to draw physical landscape and host relationships to these particular agents.

1.7.1 Predators and parasitoids

The role that avian predators play in the control of *Lambdina* populations is not well documented. Otvos *et al.* (1973) reported that numerous species of birds have been reported to prey upon *Lambdina fiscellaria fiscellaria* in Newfoundland. However, parasitoids play a more important role in the population dynamics.

Micro-parasitoids such as nematodes and protozoans are a potential source of pressure upon the looper population. *L. fiscellaria fiscellaria* have been shown to be susceptible to the effects of the nematode *Heterorhabditis heliothidis*, and larvae were killed via this predator at a laboratory-controlled temperature of 24°C (Finney and Bennet, 1984). *Microsporidia* protozoans also have been found in looper populations within British Columbia (Morris, 1983). In New Brunswick, a gregarine protozoan parasite has been found to occur frequently in *L. fiscellaria fiscellaria,* infesting 50% of larvae in 1993 and 80% in 1994, in sufficient abundance as to cause blockage of food within the gut (Lucarotti *et al.*, 1994). Smirnoff discovered the presence of a *Nosema sp.* protozoan reduced the WHL population in Quebec (Smirnoff and Jobin., 1973).

Parasitoids of WHL, attack the egg, larval and pupal stages and have the ability to affect population levels of this insect. Parasitism by dipterans is more prevalent in the larval stages. Parasitism by hymenopterans is most common in the egg and pupal stages (Turnquist, 1991). Egg parasitism by *Telenomus sp.(dalmani* Ratz.) and *Trichogramma sp.* (*minitum* Riley), has been found in 30% to 80% of eggs laid, and this was thought to cause the decline of the 1947 outbreak of WHL in British Columbia (Campbell, 1946a; Campbell 1948b). During 1984, in the Pacific region of British Columbia, it was suspected that egg parasitism was the major reason that WHL population numbers did not increase. Egg parasitism was estimated at 38%. Parasitism may work in conjunction with other factors and need not be very high to have an impact on the decline of an outbreak (Van Sickle, 1994). During the 1990-1995 outbreak in BC, egg parasitism sampled in 1993 at 34 locations averaged 21%; the outbreak declined in 1994. Conversely, over many outbreaks in the Vancouver, Nelson and Kamloops areas, the decline of the outbreaks can be attributed to egg parasitism that reached as high as 80% in some cases (Van Sickle, 1995).

Larval parasitism occurs when an egg is oviposited by a parasitoid within the WHL larvae or upon the cuticle. As incubation and hatching occurs, the parasitoid larvae feeds upon its living host and eventually emerges having killed the host. Tachinids are most often implicated with this type of parasitism in larval *Lambdina sp.*. Principal larval parasitoids include: *Winthemia occidentis* (Reinhard), *Zenilla virilis,* and *Chaetophlepsis nasellensis* (Reinhard), (Hopping *et al.*, 1938; Medley and Carolin., 1977; Otvos *et al.*, 1973; Wyatt, 1946).

Pupal parasitism also occurs commonly, as the looper is in a vulnerable life stage at this time. A wide variety of ichneumonid parasitoids attack the WHL, including: *Apechthis ontario* (Cress.), *Apechthis pacificus, Itoplectis conquisitor, Itoplectis quadricingulata* (Prov.), *Pimpla pedalis, Pimpla aquilona, Aoplus cestus* (Cresson), and *Aoplus velox occidentalis* (Harr.)(Bradley, 1993). During field collections and subsequent laboratory rearings in Alaska, the most abundant ichneumonid parasitoids were the *Aoplus velox occidentalis* (86%), *Pimpla pedalis* (5%), *Pimpla aquilona* (3%), and *Apechthis ontario* (3%) (Torgersen, 1971). These insects may yield valuable clues as to why outbreaks occur where they do, despite the fact that the looper is endemic in many areas that have never displayed outbreak-level defoliation.

Cooler, damper climatic conditions may be one factor regulating parasitoid populations. Although such conditions may still be within the tolerances of many of the regulating parasitoids, they may affect parasitoid populations in non-lethal ways. Humidity has been shown to be a significant factor in oviposition. Experiments on females of both *Apechthis ontario* and *Itoplectis quadricingulata* using *Galleria mellonella* (L.) as a host, have shown significantly fewer progeny at 100% relative humidity than at 62% relative humidity (Ryan, 1974). This serves as an example of how climatic conditions may indirectly be responsible for allowing WHL populations to rise in humid areas as compared to regions that are drier and more favourable to the parasitoids.

1.7.2 Pathogens

Looper populations are known to be adversely affected by fungal, viral and bacterial agents. In accordance with classical population dynamics there have been several models describing the host-pathogen relationship in forest insects (Andreson and May, 1980). Generally as the host density increases, so does the rate of spread of the pathogens among the host species.

1.7.2.i Fungi

Fungal pathogens have been attributed to the partial decline of looper populations (Van Sickle, 1994). *Entomophthora* sp. appear to be the primary fungal pathogens of *Lambdina fiscellaria lugubrosa,* particularly *E. egressa* and are cited as being an important natural control factor of this insect (Turnquist, 1991). In the eastern hemlock looper, *E. sphaerosperma, E. conidia, Entomophaga aulicae and Aureobasidium pullulans* are common in the larvae of collapsing populations, causing as much as 28% of the larval mortality (Otvos *et al.*, 1973; Hudak *et al.,* 1988; Hartling and Carter, 1993). During the 1966-1970 outbreak in Newfoundland, two fungal species, *E. sphaerosperma and E. egressa* were credited with the collapse of the outbreak. Tests conducted on larvae of *L. fiscellaria fiscellaria* injected with protoplasts of *Entomophaga aulicae* found that higher relative humidity and temperatures produced the shortest time until larval death and the highest conidia production (McDonald and Nolan, 1995).

1.7.2.ii Viruses

Nuclear polyhedrosis virus (NPV) causes significant mortality in both *L. fiscellaria fiscellaria* and *L. fiscellaria lugubrosa* populations, and is noted as a major factor in reducing populations during outbreaks (Cunningham, 1970).

Resistance to NPV does not increase in the 2-4th instars. The efficacy and period from ingestion to death varies and is affected by a number of factors including: the concentration of the virus, temperature at which it is applied and ingested, as well the virus' exposure to ultraviolet radiation. (Cunningham, 1970; Morris, 1971) Since temperature is a factor in the rate and degree of mortality as well as the percentage of mortality, viral success in warmer regions may actually control looper population development and cause greater mortality than in colder regions.

1.7.2.iii Bacteria

Bacteria do not appear to play as large a role in *L. fiscellaria* mortality as other forms of natural control. There are numerous forms of natural soil-dwelling *Bacillus* species, in particular *Bacillus thuringiensis* (*B.t.*) that have been implicated in looper larvae mortality, (Elgee, 1963; Heimpel and Angu, 1959; Krieg and Langenbruch, 1981). Although we are interested in natural control factors that may influence WHL distribution, *Bt* has been successfully introduced as a control method, and this topic will further be explored in the Discussion section of this thesis.

1.8 Abiotic influence-weather

Although this thesis does not aim to explain temporal patterns of WHL outbreaks, it would be remiss not to examine some of the weather patterns that may influence the activity of looper population in addition to the climatic characteristics. It is important to distinguish between the climatic influence that governs the spatial distribution of WHL populations and outbreaks, and weather that may partially contribute to the temporal patterns of population variation. Climate may play a large role in determining where core looper populations occur. Climate obviously has direct links with host and WHL distributions; however short-term fluctuations in weather conditions may provide the catalyst for a rise or a collapse in these locations. The following literature is based upon observations of populations and the weather that precedes the onset of an outbreak and the collapse of the outbreak.

Several aspects of weather may affect the looper's ability to overwinter, initiate feeding, pupate and mate. Otvos (1977) noted that outbreaks of EHL were preceded one to two years by higher than normal temperatures and lower than normal precipitation.

Overwintering mortality is certain to play a role in the population dynamics of the WHL. In samples of EHL, overwintering mortality in Newfoundland averaged 45%, 58%, 65% and 2% (from various years of sampling) when comparing eggs collected and reared in the fall against those collected and reared in the following spring. The degree of egg mortality was inversely related to the deviations from long-term normal winter temperatures (Otvos, 1977).

Larval development is adversely affected by cool wet weather in that hatching success, development and survival are hindered (Otvos, 1977; VanSickle, 1995).

Moth flight and mating are known to be affected by adverse weather conditions. Heavy rains affect flight, as do high winds (Jardine, 1969). It has been noted that the amount of rain in September is negatively correlated with WHL outbreaks in B.C. (Thompson, 1952).

Possibly one of the most important features of a population's ability to increase is larval survival rate. Synchronizing larval hatching with budflush is crucial to the WHL's ability to initiate feeding (Carroll, 1999). Baltensweiler (1993) noted that cooler temperatures tend to synchronize egg hatch so that it occurs over a shorter period. Warmer temperatures spread the hatching out. This may have the effect of de-stabilizing the synchrony with bud-flush in warmer areas, or affecting survival rates of emerging larvae.

1.9 Damage and socio-economic impact of defoliation

1.9.1 Overview of damage

Defoliator outbreaks have the potential to cause serious socio-economic losses to affected regions. The most apparent impact is upon timber supply and wood quality, although there are also tertiary effects on tourism and recreation. There are costs associated with disrupted timber management plans, a sudden over-abundance of a particular timber type derived from the salvage operations and undesirable succession from a management perspective. Non-lethal growth reduction may incur costs for years to come after an outbreak (Bowers, 1993). From a forest-management perspective, there is also the longer-term implication of the effects of defoliation on the preservation of biodiversity.

Losses from defoliators make up a substantial portion of the total annual unsalvaged timber, particularly in outbreak years. During the height of the looper outbreak and subsequent years, defoliator losses made up to 58% of the total annual unsalvaged loss, or approximately three times more than losses from fire; the next closest agent of disturbance (Table 2), (BCMOF, Kamloops TSA, Timber Supply Review, 1998). Losses on mid-seral stands could also be significant because although the trees are not yet merchantible, there can be large effects on yield at harvest. The extent of damage from WHL highlights the need to identify potential areas of defoliation in order to estimate possible mortality and establish a proactive management approach to minimizing the amount of unsalvageable loss.

Table 2. Annual unsalvaged loss in the Kamloops Region based upon the 1998 timber supply review.

Cause of Loss	**Annual unsalvaged Loss (m^3/year)***
Bark Beetle	3 900
Defoliators	***36 130***
Windthrow	9 250
Fire	12 210
Misc. (flood, etc)	1 100
Total	62 590

* based upon (1993-1998) on operable FC-1 lands

It is apparent that after defoliation occurs, trees may die as a direct result of defoliation or as a result of their weakened state that allows them to succumb to secondary pests and/or pathogens. Mortality can occur up to 4 years after a defoliation event (Humphreys, 1993). It is quite often that the mortality accelerates due to the increased activity of secondary insects. Some common secondary pests are the borers, Cerambycidae such as *Tetropodium velutinum* as well as the bark beetles, *Pesudohylensinus tsugae, Dendroctonus pseudotsugae, Dendroctonus rufipennis* (Kirby), *Dendroctonus obesus, Dryocetes septentrionis* and *Ips concinnus* (Kinghorn, 1954, Englehardt 1953). Mortality is greatest in pure hemlock stands compared to mixed stands (Kinghorn, 1954).

The percentage of stand mortality has been predicted from defoliation at the height of mortality. Kulman calculated that 80-90 % defoliation caused 65-78% mortality, whereas only 10-25% mortality occurred when stands were defoliated 50-75%. In Washington, after 70% defoliation, over 60% of the hemlock trees died within three years (Kulman, 1971). More recently and with greater precision, the percentage of stand mortality was calculated by Alfaro *et al.* (1999) as a quadratic fit of the maximum percentage of defoliation in the stand at the height of the outbreak. Combined with a hazard-rating or a prediction of the locations of outbreaks, this information can be very valuable in projecting losses from defoliation events and aiding in the overall planning and management of timber supply areas.

1.9.2 Biodiversity issues

Large looper-killed areas may have historically been large scale natural disturbance events, similar to forest fires and other landscape disturbances. WHL-induced mortality can be viewed as one of a number of the renewal processes in a forest, and may have actually served to increase the level of biodiversity by increasing the complexity of the landscape. It would have served as a successional agent to produce a variety of seral stages.

Competing interests from logging, recreation and various other uses now requires some form of forest management in order to balance these various conflicting objectives. Large scale disturbance events can disrupt this balance and are generally viewed as damaging forces that require mitigation. These interior temperate rainforests tend to experience large-scale disturbances somewhat infrequently, rather they experience smaller-scale disturbances that create more of a fine-grained mosaic (Lertzman et. al. 1997).

The *Biodiversity Guidebook,* a portion of the *Forest Practices Code of British Columbia Act,* gives guidelines for the retention of old growth forest. The ICH wk, vk and mw biogeoclimatic subzones in which the WHL occurs most severely, are primarily natural disturbance type-1 (NTD1) and type-2 (NTD2). NTD1 are rare stand-initiating disturbances with an average return time of 250 years, or NDT2 are infrequent stand-initiating disturbances with an average return time of 200 years (BCMOF, 1995). Table (3) shows the percentage of forested area that is recommended to be left in order to maintain the BCMOF objectives of preserving biodiversity.

Table 3. Recommended seral stage distribution under two natural disturbance types within the ICH wk, vk, and mw sub-zones.*

Seral Stage	**Early (< 40 years)**			**Mature (> 100 years) + Old**			**Old (> 250 years)**		
Priority	Low	Intermediate	High	Low	Intermediate	High	Low	Intermediate	High
Natural Disturbance Type-1	N/A	<30	<23	>17	>34	>51	>13	>13	>19
Natural Disturbance Type-2	N/A	<36	<27	>15	>31	>46	>9	>9	>13

* Adapted from Biodiversity Guidebook, Forest Practices Code of British Columbia Act, 1995.

One obstacle forestry managers face in maintaining biodiversity targets is that infestations such as WHL attacks can devastate large portions of old growth. If the damaged area lies within a region that has been retained to meet biodiversity requirements, then it is a threat to these objectives within the context of a managed landscape. Defining areas at risk for attack by WHL may help to anticipate shortfalls in conserved seral stages and to alter harvesting schedules to accommodate the needs of salvage operations. Rather than having rigid harvesting plans, perhaps larger portions could be set aside as 'grey' areas which would allow a more flexible approach to managing the landscape to buffer the effect of widespread looper damage. Also it could incorporate looper-affected stands (i.e. structured stands) into the landscape mosaic.

2 Materials and methods

2.1 Definition of sample groups or outcomes

It is important that the groups we are exploring are unambiguously defined. They are as follows:

- AU- adjacent undefoliated. This group contains stands that are located within the ICH-wk, vk, or mw biogeoclimatic sub-zones and within the same continuous valley as where a defoliation event has occurred but were not themselves defoliated in the 1990-1995 outbreak. Identifying stands belonging to this group are of interest delineates stands not yet affected within a valley that is at risk. Knowing the characteristics of this group may aid in reducing the overall risk of attack upon a landscape.

- L- Loci of defoliation. This group contains stands where defoliation has occurred within the first and second year of the 1990-1995 outbreak. Identifying potential loci is an important step in the placement of sentinel traps and initiating efficient site-specific damage control.

- D- Subsequent defoliation. This group contains stands that were defoliated after the first two years of the 1990-1995 outbreak. These are all other defoliated stands except the loci.

- UV- Undefoliated Stands in Undefoliated Valleys. This group contains stands that are located within the ICH-wk, vk, or mw biogeoclimatic zones, but are not considered at risk by virtue of never having been defoliated within the 1960-1995 aerial survey period. If large areas such as entire watersheds can be excluded by virtue of possessing little or no risk, it will result in better allocation of resources both during monitoring and conducting control operation during an outbreak.

The scale at which an accurate hazard-rating system can be constructed is a critical component of this study. The dataset was classified according to four potential outcomes for a stand within the ICH biogeoclimatic subzones that are considered at risk. The data are stratified first at the watershed level and then at the stand level. Figure 5 is a conceptual model of how the sampling units (stand centroids) are organized by group.

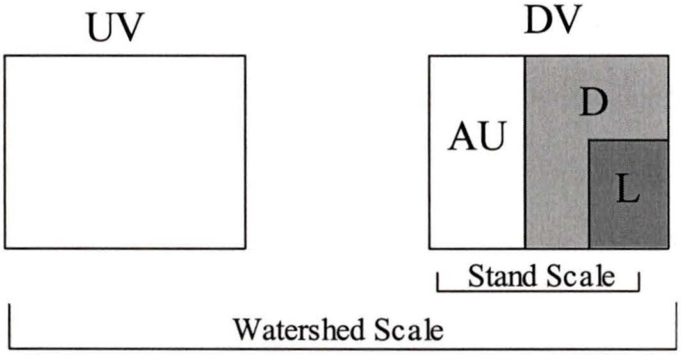


Figure 5. Conceptual model of sample groups

In grey are the defoliated groups. The watershed scale groups are undefoliated valleys (UV), and defoliated Valleys (DV) . Within DV are adjacent but undefoliated stands (AU), loci of outbreak (L) and defoliated stands (D).

It is important to determine the scale at which defoliation susceptibility can be detected. Examining the differences between defoliated loci and undefoliated adjacent stands is critical to assessing the potential for future defoliation within never-defoliated stands. We wish to find out if unaffected stands exhibit definitive characteristics that render them less vulnerable to attack, or if it is simply a matter of stochastic chance that they remained undefoliated before the *Lambdina* population collapsed.

2.2 Two step hazard rating method

A two-step method is necessary in order to accomplish the process of rating forest stands for the risk or hazard of defoliation by WHL.

2.2.1 Step One: Distinguishing between valleys at risk for defoliation

In the first step, defoliated valleys (DV) containing the loci of defoliation (L), subsequently defoliated (D) and adjacent undefoliated stands (AU), were distinguished from UV where defoliation had not occurred.

Coarse scale data, including climate surface data, and elevation were utilized within this first step in conjunction with forest cover data. Climate variables were explored with specific reference to their influence on the looper's life cycle, the role that they play in with regards to epizootic conditions and/or upon parasitoid abundance.

A dichotomous answer emerged from this analysis that determines if the valley in question is at risk or not. If the valley was type UV, then the process has shown the stands within it to have a low hazard and no further analysis was conducted. If the valley was determined to be type DV, then a second stage of discriminant analysis was performed.

2.2.2 Step Two: Distinguishing between stands in valleys at risk for defoliation.

If the valley was determined to be at risk, (i.e. it contains adjacent undefoliated stands, loci, and/or subsequent defoliation), the second discriminant analysis was performed upon the site using forest cover data in order to arrive at a function that best placed the site into an appropriate category.

2.3 Overview of rationale and study area

The location for this study is the southern-interior of British Columbia, Canada (Figure 6). This includes portions of the forest-regions of Prince George, Kamloops, Cariboo and Nelson.

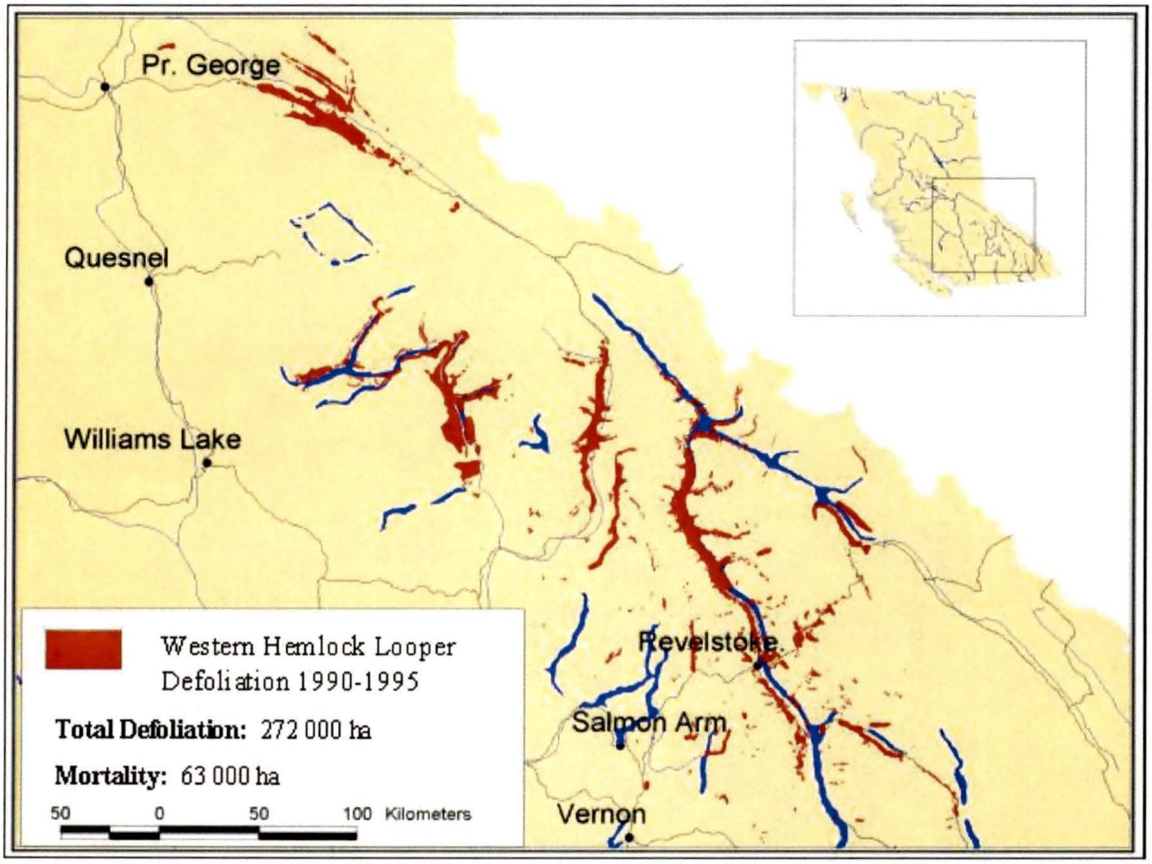


Figure 6. Location of study area and most recent outbreak (1990-1995).

Although outbreaks have taken place both in the coastal and interior regions of the Province, the interior was chosen for the following three reasons. First, the last outbreak (1990-1995) was the most extensive in the Province's history, and therefore provides the best coverage. Second, since defoliation has been recorded in BC, markedly different outbreak characteristics have been noted between coastal regions, (less extensive outbreaks), and the interior where outbreaks tend to occur over large areas and thus are more of a concern.

Third, an essential reason for choosing this time period (1990-1995) is due to the sampling constraints on available data. The available forest cover data is dated 1993, and is representative of 1990-1995 outbreak. This is an important component of the study as forest cover information and the structure of forests changes over time due to natural growth, disturbance etc. In order to obtain forest cover characteristics representative of the conditions during an outbreak, we are constrained to one particular period in time. This should be considered when interpreting the final results, however, since outbreaks are observed to take place within similar locations over time and lie primarily within the Interior-Cedar Hemlock biogeoclimatic zones, the use of the largest and most extensive outbreak recorded in history should be acceptable for our purposes.

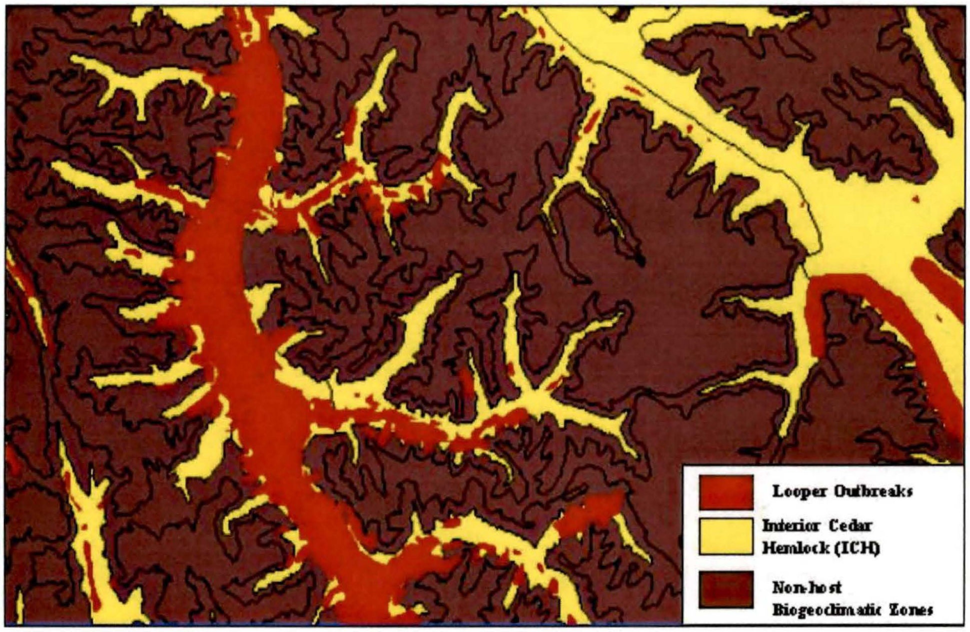


Figure 7. Western hemlock looper outbreaks compared to biogeoclimatic zones

Although the study area initially appears to be quite broad, in reality the focus was confined to areas within the Interior Cedar-Hemlock biogeoclimatic zones. Figure 7 shows western hemlock looper outbreaks in relation to biogeoclimatic zones. It is obvious that the outbreaks follow host-species distribution quite closely. In fact, over 90% of the total area of the last outbreak, (1990-1995) occurred within three subzones of the Interior-Cedar Hemlock: very wet and cool (ICH-vk1), wet and cool (ICH-wk), and moist warm (ICH-mw). When comparing the entire defoliation history of WHL from 1911 to 1995, 80% of defoliated area occurs within these three subzones, even when considering the coastal outbreaks (Appendix 1.) Outbreaks have occurred in other biogeoclimatic zones, particularly ESSF (Englemann Spruce- Subalpine Fir) and SBS (Sub-Boreal Spruce). This may be due to a number of factors including transition zones between host and non-host biogeoclimatic zones, mapping analysis error introduced by differences of scale and precision between defoliation sketch maps, and the Biogeoclimatic information, and spill-over from intensively defoliated host zones into adjacent non-host stands. For this reason, stands within the main three biogeoclimatic sub-zones are primarily considered at risk and will be the focus of both the watershed and the stand-level analysis.

2.4 Hardware and software

An extensive Geographic Information System (GIS) is already in place at the Pacific Forestry Centre. Most GIS analysis has occurred using the Environmental System Research Institute's (E.S.R.I.) software program ARC/INFO, (version 7.0). Display and map production was performed using ESRI's Arcview (version 3.1). Forest Information and Planning File data resides in an ORACLE relational database. Statistical analyses were performed using both SAS and SPSS software. Graphs were plotted with SPSS. Spatial autocorrelation analysis utilized GS+. With the exception of SPSS and GSPLUS, the software resides on a UNIX network operating on a SUN Sparcstation platform. Data is stored on a RAID system disk. The UNIX system is accessed via X-emulation software: Hummingbird's EXCEED. This runs via a Windows 98 platform, on a Digital PIII 900Mhz personal computer. Hard copy maps were plotted via HP DesignJet 755CM plotter.

2.5 Digital data

As this project is an exploratory exercise aimed at utilizing existing data to characterize risk, one of the most important steps has been to research and collect accurate digital data. It is also important to recognize its limitations. Digital data, and in particular, electronic mapping information have the potential to contain many sources of error: sampling error, equipment calibration, transformation error, classification error, machine precision, digitizing error, errors of scale, compounding, etc. Some of the main obstacles in obtaining data for this project were: data at differing scales and formats, the precision of data, lack of existing data, and limited temporal coverage. The precision of data has been noted subject to its availability, the accuracy of such data is a more difficult task to determine. Steps have been taken to ensure that the degree of error is minimized either through sample size or through various sampling techniques.

The spatial data were converted from their original formats into Arc/Info and transformed into a Lambert projection; the standard utilized for BC-wide data by the

Pacific Forestry Centre. All projections have certain compromises inherent when projecting a three dimensional sphere into two planes. Considering the extensive map overlaying and as Lambert preserves the shape of a map there is less misclassification error introduced with the use of Lambert than with other projections, even though area may be distorted. See Appendix 2 for the Lambert projection parameters used in this study.

2.5.1 Data summary

Table (4) below gives an overview of all the data used in this study. Individual variables were derived using these data.

Table 4. General Data Summary

Data	Scale	Source
Western Hemlock Looper Defoliation data (1911-1995)	1:50 000 to 1:100 000	FIDS-Canadian Forest Service 1996
Forest Inventory Planning Files (attribute) /Forest-Cover (FC-1) (spatial)	1: 20 000	BCMOF 1996
Biogeoclimatic Zones of British Columbia.	1:50 000	BC Ministry of Environment Lands and Parks
Watershed Atlas of British Columbia	1:50 000	BC Ministry of Environment Lands and Parks
PRISM Climate Surface	4000 metres	Oregon State University
Canadian Digital Elevation Data (CDED)	100 metres	Centre for Topographical Information (CTI) and Canadian Forest Service (CFS)

2.5.2 Western hemlock looper defoliation maps

The digital defoliation maps of WHL form the basis for determining which forest cover stands were either defoliated or undefoliated in the three host biogeoclimatic subzones of interest. WHL defoliation data was recorded during aerial surveys in late summer after insect feeding was completed. Survey from fixed-wing aircraft was accomplished by sketching defoliated areas onto 1:100 000 or 1:250 000 scale maps, with an estimated 200-300 m positional accuracy for the defoliated polygons (Bob Erickson NRCAN-CFS-Pacific, personal communication). Severity classes were assigned to the defoliated polygons based upon the amount of observable damage (Table 5).

Table 5. Defoliation Severity Classes

Severity	Class	Foliage Loss
Light	1	<25%
Moderate	2	25-65%
Severe	3	>65%
Dead	4	N/A

(Adapted from Parfett et al., 1995)

The FIDS rangers digitized and classified by defoliation severity, the annual aerial survey maps of defoliation since 1985. Sketch maps from regular aerial surveys starting in the 1960s were also digitized. In addition, 1911-1940s verbal reports were interpreted and digitized, although the accuracy of this early data is unknown. It appears that the data between 1930s and 1950s were a mix of verbal reports and non-systematic aerial surveys (Parfett *et al.* 1995). The construction of this database was part of a project prior to this thesis.

The defoliation data for the purposes of this study were treated as simple presence or absence data rather than by defoliation severity. If there are apparent trends in the types of stands being defoliated, they should be detectable using only presence or absence.

2.5.3 Forest Information Planning Files and Forest Cover Spatial Data (FIP/FC-1)

The forest cover information forms the primary set of variables of interest. As the most detailed and large-scale data we are looking at, the data contains an extensive array of biological information regarding forest stands. While not as accurate as ground plots, the volume of information and the aeral coverage provides a unique opportunity to examine the broader trends in forest stands affected by western hemlock looper defoliation. Sampling and collection inaccuracy and precision is offset by the ability to sample a large number of stands. In addition, forest cover data is a primary source of information that forestry managers use to make decisions. Other advantages are that this data is widely available thus new predictive models will not require extensive additional data collection.

Forest inventory data collection and interpretation is a very lengthy and complex process, with a stringent set of quality control standards in place. The forest-stand variables are drawn from air photo interpretation and ground-truthing of this data. It is important to understand that although air photo interpreters are well trained and certified within British Columbia, interpretation may still vary somewhat among interpreters as well as between transfer methods. The air photo data is collected in the following fashion.

Forest and non-forest types (polygons) are stratified through the interpretation of mid-scale aerial photographs (1:20 000 approx.), supported by field information. Some of the data are actually interpreted from 1: 5 000, 1:10 000, 1:15 800 , or 1: 31 680 . The resultant forest cover types are then transferred from the photographs directly into the digital design file using one of the following methods:

1. First order plotter (Kail or epidiascope).
2. Monorestitution.
3. Ortho photo transfer.

Appendix 3 outlines the standards for accuracy and precision where available, for the main forest cover variables utilized in this study. A more detailed explanation of the specific variables used takes place in the statistical processes section of the methods.

2.5.4 Biogeoclimatic Zones

The Biogeoclimatic Ecosystem Classification (BEC) of British Columbia was pioneered and constructed by the late landscape ecologist, Vladimir Krajina and his students in the 1960s and early 70s. From 1976 onwards, the BC Ministry of Forests began data collection and classification for the entire province; a process that continues to be updated to this day (BCMOF, 2001). Although not as detailed a dataset as FIP/FC-1 data, they represent one of the single most useful tools in classifying and managing landscapes. Biogeoclimatic data are stratified into zones, subzones, and variants (phases). There are 14 zones in British Columbia, representing the climax or late seral stage species found within a large homogenous area exhibiting a specific macroclimate.

Sub-zones are delineated by using indicator species and plant associations or occurrences on a site. Subzones may include differences in proportion and vigour of certain indicator species, or variations in successional development, and are indicative of the variations in climate (cooler, warmer, wetter, drier) within that particular zone. Variants (Phases) are subzones further stratified by unique topography, aspect, or climatic pockets and are usually referred to as a specific location. The exact methodology is a complex hierarchy of plant orders, plant alliances, associations and sub-associations (Krajina, 1965; Pojar *et al.* 1987). For the purposes of this study, BEC delineates the main location of WHL's primary host species, the western hemlock (*Tsuga heterophylla,* (Raf.) Sarg.), within the Interior Cedar-Hemlock biogeoclimatic zone.

2.5.5 British Columbia Watershed Atlas

Although the actual watersheds did not provide any attribute data used in the analysis, it was necessary to have a standardized method of delineating valley boundaries for the first stage of analysis. Before forest managers begin an evaluation of stands that are at risk within their regions, it is important to know if a particular valley (watershed) is at risk. In order to perform a valid statistical analysis, the sampling units should follow some standardized rule and be both repeatable and identifiable. The BC Ministry of Environment Lands and Parks, Fisheries Branch (circa 1998), has produced a watershed atlas of British Columbia. Watersheds are delineated via a system of modeling that takes into consideration the lands (to the height of land) drained by a specific size stream or river including all feeder channels based upon aquatic features from NTS 1:50 000 scale maps. The form of these data is simple for the purposes of this study. In order to delineate the concept of a 'valley at risk', a valley was defined as being a third order watershed or higher. Such a valley would include all class three streams and lower to the confluence of a higher order stream. It should be noted that this is a hierarchical system whereby third-order watersheds contain second-order watersheds, and fourth-orders contain third (BC Ministry of Environment, Lands and Parks and Spatial Vision Consulting, 1998). Higher order watersheds were used if these watersheds did not contain any third order watersheds. This would include large river valleys such as the Thompson River or Upper Shushwap.

Third order watersheds formed continuous polygon coverage for all areas of interest. For a watershed to be considered a 'defoliated' watershed, only one forest stand had to be defoliated within that polygon.

Very large (usually 4th order) watersheds presented somewhat of a problem. There was one exception in the Upper Shushwap watershed, where a very small patch of defoliation (142 ha) occurred at the northern boundary of the watershed, caused the entire length of the watershed to be classified as defoliated. Considering that the defoliation occurred within the first 5 km of a 135km long watershed, and on the shared boundary of a heavily defoliated watershed, it seemed improper to classify the entire Upper Shushwap as falling into the defoliated valley category. It was judged that as a whole, the watershed had not experienced defoliation. Other than this particular case, undefoliated watersheds were defined as not having had defoliation between 1911-1995. Defoliated watersheds were those that had experienced defoliation between 1911-1995.

It was initially planned that some watersheds prior to the start of the 1990 outbreak would be excluded entirely from the analysis on the grounds that there was no forest cover data prior to 1990. However, since the last outbreak was so extensive, all watersheds that were positive for defoliation were included in the analysis.

2.5.6 PRISM climatic data

From looking at the defoliation data, it was apparent that large segments of the ICH biogeoclimatic zones do not get defoliated; a consistent trend over time. Defoliated areas occurs within the same general area, although not consistently in the same places. As these defoliation patterns occur over a large area, utilizing a climate model appears to be the best way to initially isolate large areas at risk for defoliation from those that are not. Climate surface data is used in conjunction with watershed data for the coarse scale assessment of watersheds that exhibit a tendency to be defoliated by WHL.

Two groups have constructed Canadian climate models. One model called ANUSPLIN has been created by D.W. McKenney of the Canadian Forestry Service in Sault Ste. Marie, Canada (Mackey *et al.* 1996).. The other model is called PRISM and was created by Christopher Daly of Oregon State University, U.S.A. (Daly *et al.* 1994, Daly *et al.* 1997). There are various viewpoints on the merits of either system.

Certain considerations were given before applying Daly's model to our situation, and there is no lack of debate over the 'right' climate model. The ANUSPLIN model which uses a trivariate smoothing spline to interpolate climate as a function of latitude, longitude and elevation appears to work very well with gentle topography, but is less suited to handle the complex terrain characteristic of the B.C. Interior (Ross Benton, NRCAN-CFS-Pacific, personal communication). Personal review of the ANUSPLIN model deemed that the accuracy of Dan McKenney's analysis would likely be useful enough for the type of surface that we would require for this hazard rating, particularly as it will soon be incorporating a function that considers large water-bodies and rain-shadows. At the time of this analysis, PRISM was a more readily available and possibly more accurate choice existed. The ability to handle complex topography, rain-shadows, as well as lake and ocean effects make this a better climate model for the needs and location of this study (Daly *et al.,* 1994; Daly *et. al.* 1997). The claim to PRISM's suitability for B.C. is that:

> "ANUSPLIN is OK for approximate maps and is appropriate for high-quality maps in data-poor areas. The main problem is that it does not vertically extrapolate relationships between a climate variable and elevation. So places like the Coast Range of BC, where there aren't any high elevation stations, you are out of luck. PRISM does extrapolate vertically. Also, PRISM accounts for coastal effects on temperature, and has a trajectory model that transports moisture inland over complex terrain to determine rainshadows and wet zones. This is not even attempted in ANUSPLIN."
>
> - Christopher Daly, e-mail communication.

The PRISM climate surface is available for all of British Columbia presently. One of the drawbacks of PRISM is that it is somewhat of a black box where proprietary algorithms input 30-year climate normals from BC weather stations and output the resulting interpolated climate surface. The resolution for these data is 2.5 minutes latitude and longitude or about 4 000 metres (*Ibid,* 1997). For this analysis, we utilized all available climate data. This was maximum, minimum and mean temperature as well as precipitation for all stands in watersheds within the ICH biogeoclimatic zones.

2.5.7 Canadian Digital Elevation Model (CDED)

The source of the elevation data used for this analysis was the CDED model derived from the National Topographic DataBase (NTDB) digital files at a scale of 1:50 000. This data has a resolution of 0.75 arc seconds in the easting and northing, or approximately 23m N by 11-16m south. Elevation and aspect of individual forest stands were calculated by taking an average of data points within the stand. This data was converted from NAD83 to NAD27 datum in accordance with the rest of our digital cartographic data sets. This data was produced by the Canadian Forest Service, Ontario region, and the Centre for Topographical Information by utilizing the Australian National University Digital Elevation Models (ANUDEM) software. It is judged as being both more accurate and precise than required for estimating stand elevations.

2.6 Mitigating sampling error

One of the main sources of potential error in using the information available involves the use of data at differing scales and precision. In particular, the inherent error associated with comparing forest cover data collected at a 1: 20 000 scale to the 1:250 000 scale defoliation sketch maps. The defoliation polygons for 1990-1995 were buffered by 300 metres both inside and outside to correct for this possible sampling error (Figure 8).

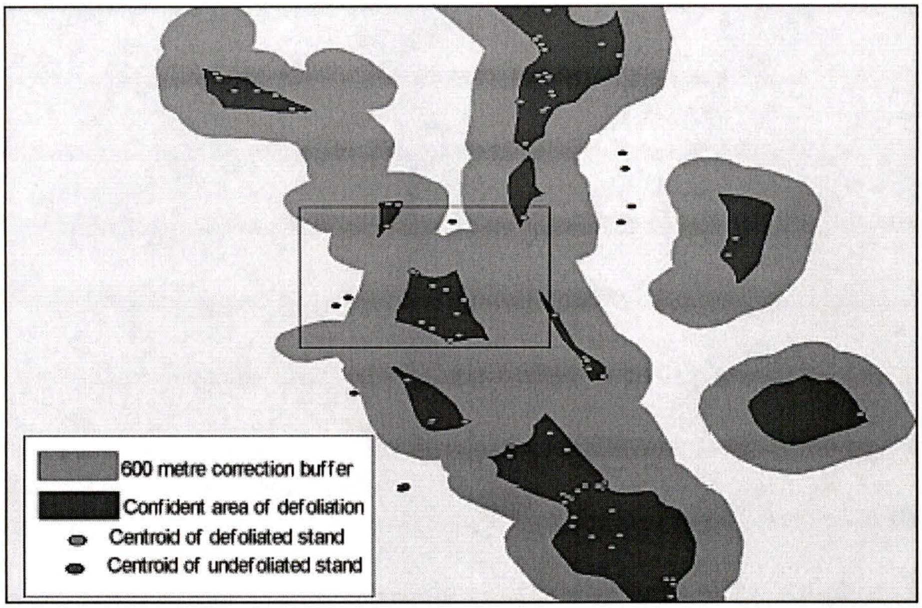


Figure 8. Defoliation area corrected for possible error.

This was the estimated region of error based upon the FIDS rangers' ground-truth efforts (Bob Erickson-CFS, personal communication 2000). This area of uncertainty was not sampled to avoid any error of providing a positive sample where there should have been a negative sample, or vice versa.

Once the potential area was compensated for by the 600 m buffer, positively defoliated stands were defined as those that had any area intersecting a historical defoliation polygon. Undefoliated stands were those that did not intersect either a defoliated polygon, nor the area of uncertainty generated by the buffering.

As the data was current as of the initiation of the outbreak, but not updated upon a yearly basis, a three to five year lag can be associated with the age status of these stands. In order to eliminate uncertainty regarding stand structure during the occurrence of the 1990-1995 outbreak, all stand age data that could not be substantiated as being undefoliated during this time, was excluded from the dataset to eliminate false negatives.

This removed all stands < 5 years old. The shortfall of this exclusion is that it does not allow the identification of defoliation risk in areas harvested and replanted within the 0-5 year range. As outbreaks have never been observed to begin in early silviculture blocks (which offer little in the way of significant amounts of foliage or protection from predation), this exclusion is not expected to present any shortcomings in the final model.

2.7 Obtaining samples of the four groups

In keeping with the two-stage hazard-rating process, it was necessary to sample in two stages. These samples were collected utilizing polygon over polygon GIS analysis known as overlaying or intersecting.

I)

i) Utilizing the BC Watershed Atlas's 3rd-order watershed boundaries, and biogeoclimatic zone data, all watersheds that lie within the biogeoclimatic zones ICH-vk, mk, and mw, were selected and considered as potential candidates for defoliation.

ii) These watersheds were overlaid with FIDS defoliation maps to classify defoliated watersheds/valleys (DV) and undefoliated watershed/valleys (UV). The actual forest stand is the sampling unit, but stratified by watershed.

II) In defoliated valleys, FIDS defoliation data were used to classify forest stands. WHL defoliation data were separated into two groups. The first contained areas defoliated in years one and two of the outbreak. Forest stands that intersected these polygons became the loci samples (L). Year one alone yielded too few stands and was unrepresentative of the entire range of outbreak. Areas defoliated in years three and beyond were called subsequently defoliated stands (D). Stands that did not meet either of these two conditions, yet still lay within the defoliated watershed and ICH biogeoclimatic zones were classified as undefoliated adjacent (UA) samples.

When the forest cover data were classified into one of the four groups, the resultant GIS polygon samples were linked to the BCMOF's Forest Information and Planning Files through an SQL connection in ORACLE . Each sampled forest stand polygon was selected for completeness and continuity of attribute data . About half of the samples were rejected due to incompleteness in the attribute database. This is primarily due to the fact that the data was outside of the FIP-file jurisdiction; that is private lands, Tree Farm Licenses (TFLs), or within parks. A visual inspection of rejected polygons did not reveal any blatant gaps in the representation of the sampled landscape; that is stand sampling was spaced fairly evenly throughout the ICH biogeoclimatic zone.

2.8 Statistical techniques

Both discriminant function analysis, and multiple logistic regression answer the similar questions. Discriminant Function Analysis seeks to parameterize a model that best puts forest stands into their correct category. Although it follows standard parametric assumption, a non-parametric equivalent is available. The advantage is that it allows the interpretation of a large number of groups, and since we are dealing with four different forest stand conditions, it is the most appropriate given the exploratory nature of this analysis.

Logistic regression works with binary data; in this case, either defoliated or undefoliated. Although multi-way comparison models can be built, this solution is less elegant and more prone to statistical error propagation than discriminant analysis. Logistical models have the advantage of not requiring the same assumptions as discriminant analysis, and estimate model parameters via a method of maximum likelihood (Bergerud, 1996).

2.8.1 Discriminant analysis statistical process

After obtaining a lengthy list of variables of interest, several analytical procedures were performed sequentially with the aims of testing each individual variable's ability to separate between groups, as well as its ability to meet the assumptions required of discriminant function analysis and Canonical Discriminant Analysis (CDA), the latter being a flavour of discriminant analysis.

1. Each variable was visually examined individually by group (UA, L, D, and UV) with a histogram for normality, and to assess its distribution.
2. Descriptive statistics were run to examine standard deviation from mean, skewness and kurtosis.
3. Variables that were not normal were transformed into their normal equivalents.
4. Correlation matrices were developed to assess the degree of multicollinearity among the variable set.

5. Moran's I test of spatial autocorrelation was performed for forest cover data based upon areal-weighted centroids.
6. Mann-Whitney U-test (non-parametric) was used to test differences between the means of defoliated and undefoliated groups.
7. The Mann-Whitney U-test was confirmed by non-parametric one-way ANOVA utilizing Kruskall-Wallace Tukey-type multiple comparison tests for differences between individual sample groups (UA, LD, SD and UV) for each variable of interest.
8. When Box's M test for equality of group covariance matrices found significant differences between groups, the covariance matrices were examined for radical differences. A ten-fold discrepancy in the absolute covariance value or a change in signs indicated a violation of the assumption of equal group covariance to the degree that the analysis may provide an inconclusive or even a false answer.
9. Two step CDA-analysis was performed:
   - i) To separate defoliated and undefoliated valleys (UV vs. DV) at the watershed scale.
   - ii) To separate (UA, LD, and SD) stands within defoliated watersheds.

Requirements for best Canonical Discriminant Function model:

In the modeling process, a large number of possible combinations of variables can be utilized to form similarly competent functions. It is necessary to have a number of criteria by which to measure the relative superiority of one combination of variables over another.

1) **Maximize Canonical Correlation.** This is a measure of the association between the function's discriminant scores and the groups. In our case, the higher the canonical correlation, the more strongly that the discriminant function distinguishes between defoliated and undefoliated tree stands. With just two groups, such as in our initial discriminant analysis that deals with distinguishing between two different valley types, the canonical correlation is the equivalent of a Pearson correlation. With more than two groups, the canonical correlation is the square root of the ratio of the between-groups sum of square to the total sum of squares, otherwise similar to the *eta* value from a one-way ANOVA (SPSS, 1999).

2) **Maximize Classification Prediction.** The coefficients of each variable derived within the CDF analysis, and the corresponding constant, are used to predict group membership for each particular case. The predicted and actual group membership values are then compared for all cases. This gives us an accuracy measurement of our models by comparing correctly classified to misclassified cases. Cross validation is an extension of this. There are two methods of accomplishing this. SPSS has a leave-one-out method of cross validation to adjust for the overly-optimistic classification that results when using classification functions derived from the same cases that they are predicting. The leave-one-out method uses discriminant functions derived from all cases except the one being classified (SPSS, 1999). An alternate, and much more robust method of cross validation uses a 'hold back' dataset not used in any calculations of the discriminant

function. This will be further discussed within the results section. Generally, this form of cross validation results in far fewer correct classifications than found in the original data or the ‘leave-one-out’ method.

Prior to attempting to construct a model of looper susceptibility for forest stands, it is important to explore some of the relationships, or the potential of relationships between defoliated stands and relevant stand characteristics. In this manner, a better model can be produced if we know, *a priori,* to the modeling process, which stand characteristics aid in distinguishing between defoliated and undefoliated stands. Figure 9 illustrates the process of identifying relevant variables to include in such a model.

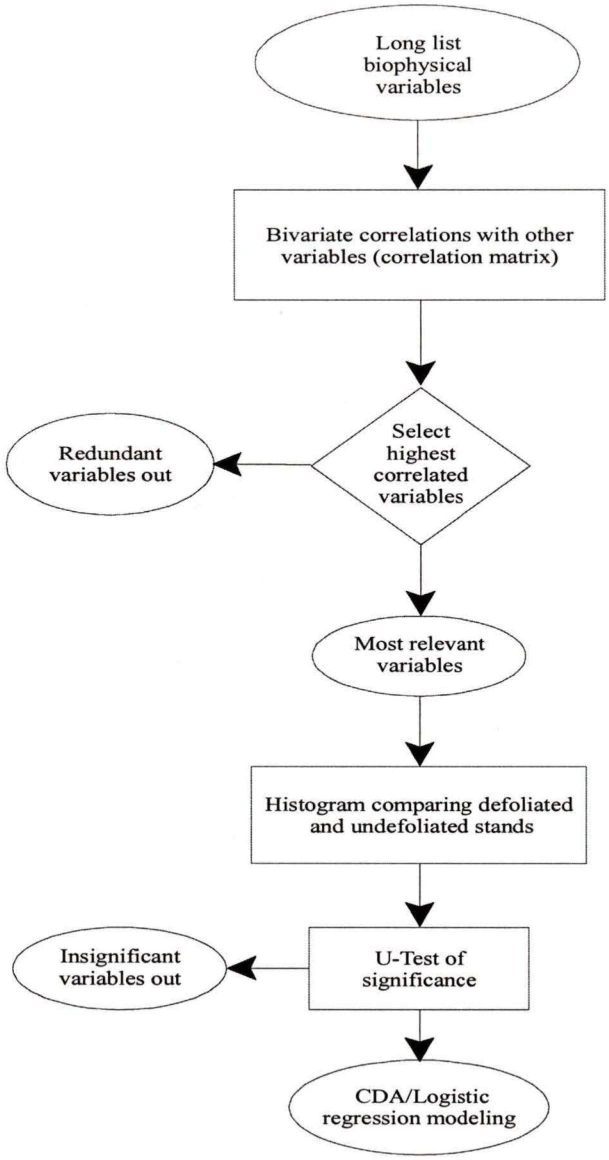


Figure 9. Variable selection process.

2.8.2 Logistic regression

Logistic regression is a form of regression that is used when the dependent variable is a dichotomy (in our case 1 = defoliated, 0 = not defoliated). Logistic regression applies maximum likelihood estimation after transforming the dependent into the natural log of the odds of the dependent occurring or not.

Basic Form of a Logistic Model:

Prob (defoliation) = $1/(1 + e^{-Z})$

Where $Z = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta n X_n$

The logistic regression model, belongs to a class of models called *generalized linear models* . These are models that, after appropriate transformation, have a linear response with a stochastic component that does not require a normal distribution (Hosmer and Lemeshow, 1989). We also must follow some standard rules applying to the logistic regression models:

- Experimental and sampling units are randomly selected from a clearly defined and appropriate population.
- Responses (counts of success) are independent.
- Values of independent variables are accurate.
- The proportion of response counts must be between 0 and 1, but not at the extremes (i.e. we can not accurately estimate a model which has 99 negative values and 1 positive value)
- Response counts are binomially distributed.
- The probability of success at a specific x-value (i.e. for a single x model), is constant.

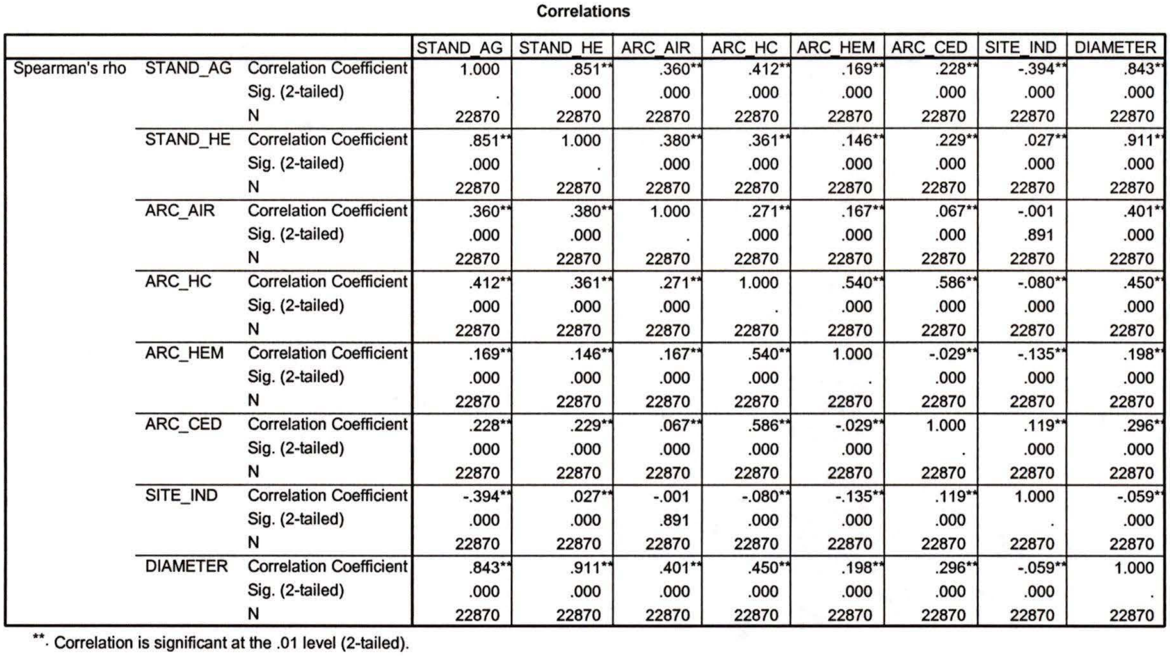
- Response counts at any x-value have binomial variance of $NY(1 - Y)$, which depends on X if Y depends on X.
- If a discrete qualitative variable is a predictor, it is encoded by dummy x variables (i.e, age class or crown class).
- Hypothesis tests rely on large sample sizes.
- Unlike ordinary least squares regression, logistic regression does NOT assume that the X variables are independent of one another. This does not relieve us of the problem of multicollinearity. Although the Y may have an accurate fit, X-variables that have a near linear relationship will cause the b-coefficients to be wildly variable.
- The coefficients (b_1), are the amount by which the expected value of Y increases when X_1 increases by a unit amount and all the other X variables are held constant. (This does not hold water if some of the X variables are functions of others (i.e. interaction terms).

(after Bergerud, 1996)

2.9 Multicollinearity

In order to ensure the orthogonality (independence) of the variables included in the model, it was necessary to generate a correlation matrix on the components that may have been of use in modeling forest stands susceptible to looper attack (Table 6). This allowed redundant variables to be identified within our list of possible parameters. Based upon the identification of highly correlated variables, decisions can be made, as to which of these variables explain the most amount of information and make biological sense.

Table 6. Correlation Matrix of Forest Stand Variables.



Correlations

			STAND_AG	STAND_HE	ARC_AIR	ARC_HC	ARC_HEM	ARC_CED	SITE_IND	DIAMETER
Spearman's rho	STAND_AG	Correlation Coefficient	1.000	.851**	.360**	.412**	.169**	.228**	-.394**	.843**
		Sig. (2-tailed)	.	.000	.000	.000	.000	.000	.000	.000
		N	22870	22870	22870	22870	22870	22870	22870	22870
	STAND_HE	Correlation Coefficient	.851**	1.000	.380**	.361**	.146**	.229**	.027**	.911**
		Sig. (2-tailed)	.000	.	.000	.000	.000	.000	.000	.000
		N	22870	22870	22870	22870	22870	22870	22870	22870
	ARC_AIR	Correlation Coefficient	.360**	.380**	1.000	.271**	.167**	.067**	-.001	.401**
		Sig. (2-tailed)	.000	.000	.	.000	.000	.000	.891	.000
		N	22870	22870	22870	22870	22870	22870	22870	22870
	ARC_HC	Correlation Coefficient	.412**	.361**	.271**	1.000	.540**	.586**	-.080**	.450**
		Sig. (2-tailed)	.000	.000	.000	.	.000	.000	.000	.000
		N	22870	22870	22870	22870	22870	22870	22870	22870
	ARC_HEM	Correlation Coefficient	.169**	.146**	.167**	.540**	1.000	-.029**	-.135**	.198**
		Sig. (2-tailed)	.000	.000	.000	.000	.	.000	.000	.000
		N	22870	22870	22870	22870	22870	22870	22870	22870
	ARC_CED	Correlation Coefficient	.228**	.229**	.067**	.586**	-.029**	1.000	.119**	.296**
		Sig. (2-tailed)	.000	.000	.000	.000	.000	.	.000	.000
		N	22870	22870	22870	22870	22870	22870	22870	22870
	SITE_IND	Correlation Coefficient	-.394**	.027**	-.001	-.080**	-.135**	.119**	1.000	-.059**
		Sig. (2-tailed)	.000	.000	.891	.000	.000	.000	.	.000
		N	22870	22870	22870	22870	22870	22870	22870	22870
	DIAMETER	Correlation Coefficient	.843**	.911**	.401**	.450**	.198**	.296**	-.059**	1.000
		Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.
		N	22870	22870	22870	22870	22870	22870	22870	22870

**. Correlation is significant at the .01 level (2-tailed).

Since we are dealing with forest cover data, it was expected that certain variables would be quite highly correlated (Table 6). As many of our variables were skewed, and required alternate forms of transformation to normalize them (were it possible at all), the decision was made to perform a Spearman's Rank Correlation upon candidate variables. Although this is less powerful than Pearson's Coefficient of Correlation, an analysis of Pearson's yielded similar results. As eight variables were tested for inclusion in the model, it is possible that some spurious correlation may exist.

It is apparent that stand age was highly positively correlated with both stand height, and stand diameter (diameter of the host species, *Tsuga heterophylla*) with rho = 0.851 and rho = 0.843, respectively (Table 6). This is certainly expected. Since cedar and hemlock content additively comprise the arc-transformed 'ARC_HC' variable, we see it to be positively correlated with both hemlock (ARC_HEM rho = 0.540) and with cedar (ARC_CED rho = 0.586). Although all other correlations were significant, other variables were only weakly correlated, or mildly correlated at rho < |0.5|. To avoid redundancy in the modeling process, only one of a pair of bi-variately correlated variables exhibiting rho > 0.5 was included in the modeling process.

The correlation matrix is a good place to begin when considering the list of possible variables to include in the exploratory stages of discriminant analysis. In particular, stand age was highly correlated with height and diameter, but the latter two made less sense to include from a biological point of view. Age tends to describe the diameter and height of stands quite well, but also can be associated with various types of understory species, canopy structure, favorable ovipositioning locations such as thick duff, Old Man's Beard (*Usnea barbata)*, bark crevices, and other features that have not been directly measured within the forest.

2.10 Spatial autocorrelation

The problem of spatial autocorrelation with respect to forest coverage and defoliation data cannot be ignored. A variable is said to be autocorrelated if its value in a specific place or time is closely correlated with its values in other places or times (Odland, 1988; Cliff and Ord, 1973). This study was attempted to account for potential problems due to spatial autocorrelation. Spatial autocorrelation is an artifact of virtually all spatial data by virtue of the fact that phenomena at one location (for example tree age), is likely related to phenomena at locations nearby.

Spatial autocorrelation can be of great aid in attempting to predict the values of locations between point sampled data (interpolation). However, it can also present some serious statistical difficulties when attempting to apply classical hypothesis testing or regression modeling. Since both require independence among observations, obviously data that is related in space or within a certain distance fails to satisfy this basic tenet. Autocorrelated observations can lead to biased estimates of the standard errors used in many hypothesis tests. Biased estimates of the standard error can drastically alter the conclusion when applying for example, Student's t-distribution for comparing means, which relies on an accurate estimation of the standard error.

In regression analysis and discriminant analysis, independent residual errors are also an issue, particularly when fitting a model to spatially related biological inventory data. If the standard error is underestimated (when there is positive spatial autocorrelation), then the significance of regression coefficients will be inflated, possibly leading to the false conclusion that certain coefficients are significantly different than zero. In addition, if residuals are autocorrelated, it means that the form of the model is incorrect and some source of variation has been omitted from the model.

There are two options available for dealing with spatial autocorrelation for these particular data: i) modeling the autocorrelation function, or ii) sampling it out of the modeled data based upon the distance (known as the lag) at which it occurs. As this study is designed with the idea that the resulting models should be fairly straightforward to implement, adding an additional autocorrelation function is not as desirable as removing the autocorrelation in the data-gathering stage. The best way to accomplish this is to alter the sampling schedule so that either only those samples at the lag beyond which autocorrelation no longer occurs are included (Figure 10), or all pairs falling within this distance are examined and one of the pair is arbitrarily dropped from the final analysis. The former technique has been successfully demonstrated in several studies using logistic multiple regression (Li *et al.,* 1997; Pereira and Itami, 1991).

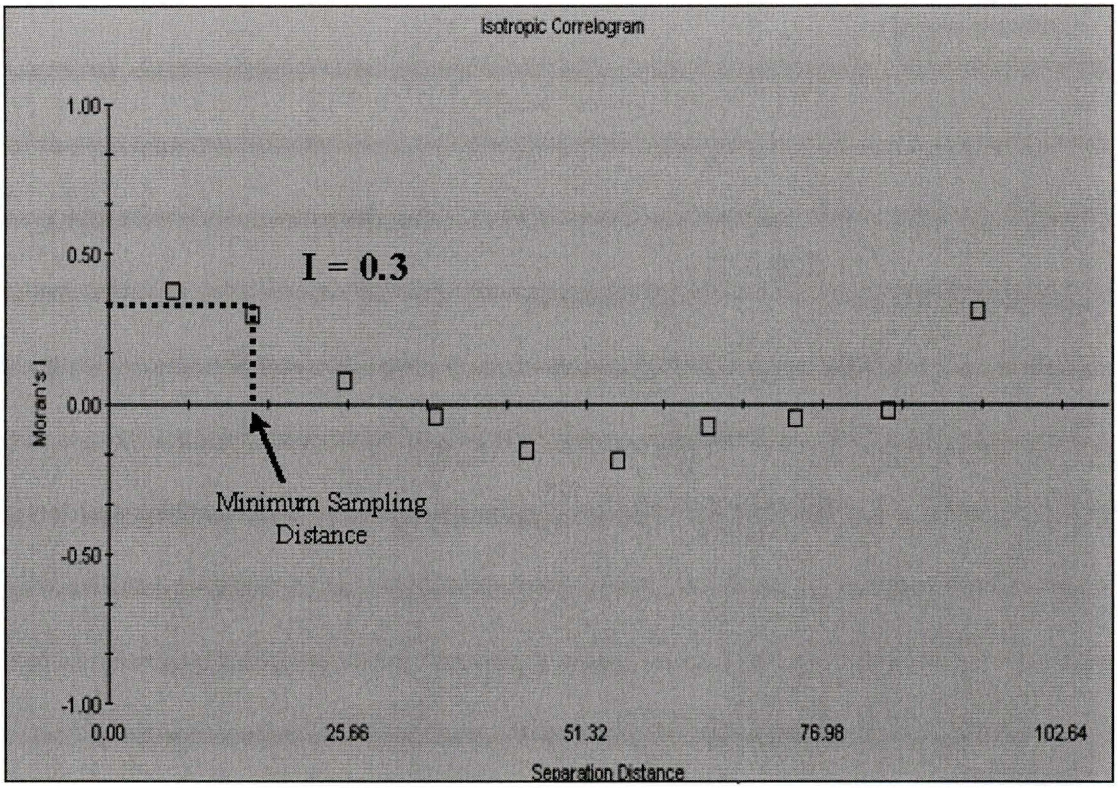


Figure 10. Sampling distance as defined by zero spatial autocorrelation

Spatial autocorrelation was tested according to Moran's Index of spatial autocorrelation, through visual assessment of variograms, and pair-wise ranking of Moran's I.

The Moran's I statistic takes the form:

Equation 1. Moran's Index of Autocorrelation (Bailey and Gatrell, 1995).

$$I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\left(\sum_{i=1}^{n} (y_i - \bar{y})^2 \right) \left(\sum \sum_{i \neq j} w_{ij} \right)}$$

I = autocorrelation for interval distance class w;

y_i = the measured sample value at point i;

y_j = the measured sample value at point i+h.

w_{ij} = proximity matrix.

Rather than examine the entire dataset, a sub-sample of 2000 forest stands were selected to test for the influence of spatial autocorrelation. As this is a fairly computationally-intensive process, 2000 stands were deemed sufficient to cover a number of lags. Each individual variable entered into our analysis was tested for the presence of spatial autocorrelation using a geostatistical package called GS+. A distance matrix was constructed using areal-weighted centroids from each forest-stand polygon. This is an accepted means of determining proximity when using areal data (Bailey and Gatrell. 1995). Several scales were examined from 5 metres to 2000 metres for lag distances, over a total of 14 500 metres), and searching for spatial anisotropy (directional influence) at 22.5° interval axis. The output for this analysis is given in Appendix 4.

Most variables exhibited a fairly low degree of spatial autocorrelation. The variable *Site Index* exhibited largest Moran's I (0.53) at a lag of 100 metres based upon 120 pairs of records (each record is paired for a certain lag distance with all other records, think of it as a giant distance matrix between all points). As the Moran's statistic is analogous to Pearson's R, it is a judgement as to what constitutes problems within the context of our analysis. Anything above I = 0.5 presents some concern, however the majority of lags for all variables exhibited a Moran's I of below 0.3. A very small number of pairs did exhibit an I-statistic in the 0.5 range for each variable. It was concluded that there was not sufficient evidence to alter the sampling schedule to correct for the presence of autocorrelation.

Initially, this was somewhat surprising, particularly since one might expect similar forest structures to exist at some scale, i.e. clustering. One of the reasons for the lack of autocorrelation may be the result of patchy harvesting and disturbance events, leaving an irregular pattern upon the landscape. Another may be that defoliated stands and undefoliated stands were lumped together. A proxy method of comparing defoliated and undefoliated watersheds may be examining the patterns of spatial autocorrelation of variables in each.

2.11 Modeling variables

The list of variables that were to be explored in the final modeling stages are listed below. The abbreviated version used in the output is listed in squared brackets

Average precipitation [AVPPT]

Average maximum temperature [AVTMAX]

Average minimum temperature [AVTMIN]

Elevation [Mean-z]

Stand age [STAND_AG]

% Hemlock content (*Tsuga heterophylla)*[HEMPRO]

% Cedar content (*Thuja plicata*) [CW]

% Hardwoods [HARDWOOD]

% Douglas-fir [FD]

% Spruce (*Picea glauca, Picea engelmanii)* [S]

% Crown closure [AIR_CROW]

Site index [SITE_IND]

2.12 Basic description of the final variables

1. **Average precipitation [AVPPT]**

   The average annual precipitation in centimetres.

2. **Average maximum temperature [AVTMAX]**

   The average maximum temperature for each month in degrees Celsius summed and divided by 12 months.

3. **Average minimum temperature [AVTMIN]**

   The average minimum temperature for each month in degrees Celsius summed and divided by 12 months.

4. **Stand age [STAND_AG]**

   This represents the mean age of the dominant tree species in the stand. A stand being identified as:

> “ a community of trees possessing uniformity in composition, age, arrangement or condition so as to be distinguishable from forest or other growth on the adjoining area, thus forming a sivicultural or management entity.” (BCMOF, 1998)

5. **% Species content**:

This represents the total amount of each species within a stand, expressed as a percentage of the total stand area.

5a. % Hemlock and cedar content [HEMCED]

Summation of western hemlock and western red cedar expressed as a percentage of the total stand area.

6. Crown closure [AIR_CROW]

The percentage of ground area covered by the vertically projected tree crown areas based upon air photo interpretation.

7. **Site index [SITE_IND]**

An expression of the growth potential of a site, based on either the site’s average stand height or the top height at a specified age, the latter being a more objective measure. This is based upon site index equations for leading individual tree species in British Columbia. Site index is calculated using the average height of top height trees of a certain species at 50 years breast height age. Top height trees are the 100/ha largest DBH suitable trees of one species. Breast height age is the number of annual growth rings in a tree at breast height (1.3 m). (BCMOF, 1998)

The accuracy of the site index for individual stands is of medium reliability. The stand's site index may have been estimated in some cases from regression results or from data of a lower sample size than noted above. These estimates of site index can vary for individual stands as a result of a variety of factors including: microsite differences in very wet and dry sites, soil strata differences (affecting soil temperature, nutrient availability, aeration etc.), aspect and elevation, (affecting snow duration, temperature, moisture conditions), as well as a latitudinal differences (BCMOF, 1998).

2.13 Comparison of defoliated vs undefoliated forest stands in ICH wk, vk, mw biogeoclimatic subzones.

In addition to removing or reducing the number of possible variables to be included in the initial model by considering the issue of redundancy in our dataset, it is reasonable to assume that if a variable is a good candidate for inclusion there will be logical population differences in the distribution and means of the variable between defoliated and undefoliated stands. Discriminant analysis and logistic regression can make use of these differences of to distinguish between defoliated and undefoliated stands. Initially, one of the ways to examine this was through a simple histogram of averages in each of the four forest-stand groups of interest: **L, D, AU,** and **UD**. This can be confirmed by a t-test or in the case of data that is inhibited by non-normal distributions that cannot be uniformly transformed, by a non-parametric Mann-Whitney U-Test.

It is of key importance to note that while some variables exhibited departures from normality that were not corrected by any transformation method, discriminant analysis is fairly robust to some departure from normality. After parametric discriminant models were run, the results were compared to a non-parametric discriminant analysis using a kernel method. The results were comparable.

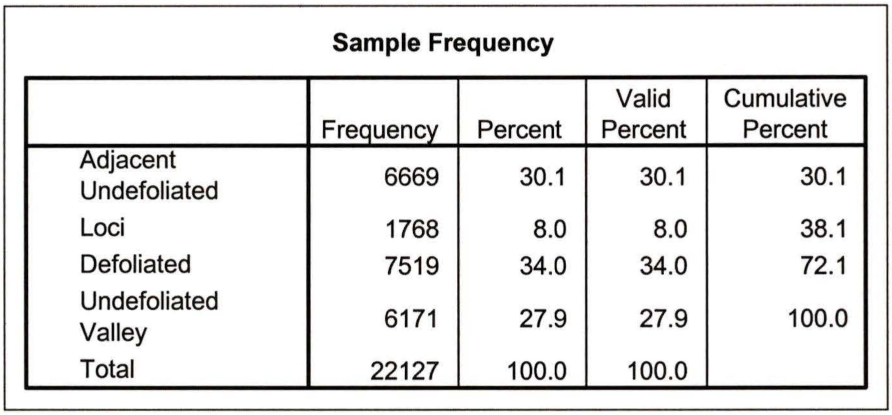
3 Results

3.1 Exploration of candidate variables

3.1.1 Sample frequencies

Overall, frequency analysis of the dataset showed that it was biased in favour of non-defoliated stands . In this case, there are about 15% more undefoliated stands than there are defoliated in the study are.

Table 7. Frequency distribution of sample types



Sample Frequency

	Frequency	Percent	Valid Percent	Cumulative Percent
Adjacent Undefoliated	6669	30.1	30.1	30.1
Loci	1768	8.0	8.0	38.1
Defoliated	7519	34.0	34.0	72.1
Undefoliated Valley	6171	27.9	27.9	100.0
Total	22127	100.0	100.0	

From Table 7, it is apparent that the sampling methods used have resulted in similar sample sized for three of four groups. Since the loci, or start of an outbreak is smaller than the scale of defoliation by the end of an outbreak, it is logical to see fewer loci stands than subsequently defoliated stands. Outbreak probability estimation using logistic regression works best when there are similar numbers of positive and negative cases (Bergerud, 1996).

3.1.2 Exploration of variable means

Initial histogram analysis among all four groups revealed differences in the distributions of our variables of interest. Age is presented as an example of the same analysis that was conducted for each forest cover variable. Not all stand variable histograms are shown in this section for the purposes of brevity and eliminating redundancy. The rest can be found in Appendix 5.

Comparisons between various age histograms suggests that the stands with defoliation tend to be more normally distributed, with an older mean age than undefoliated stands (Figure 11). This trend is clearest in loci stands.

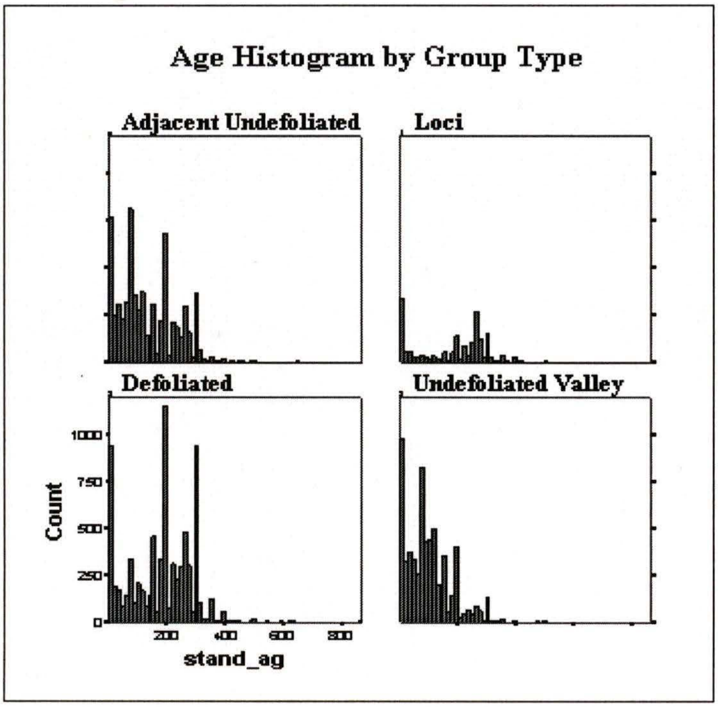


Figure 11. Age histogram comparison of the four group types

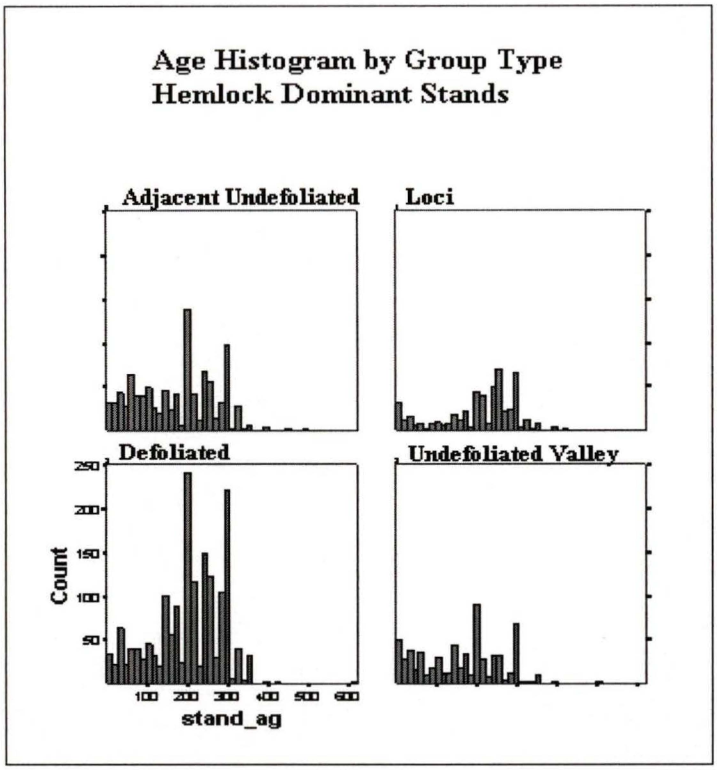


Figure 12. Age histogram of the four groups in only hemlock-dominant stands

.

As stand age refers to the mean age of the dominant species in the stand, it is not necessarily representative of the host species, hemlock (*Tsuga heterophylla*). To visually assess how this compares to the age distribution of the preferred host species, and decide whether this will have an impact on our data, we sampled for cases where hemlock was listed as the primary species. This means that it was either the most dominant in terms of ground area, or equal in area to other species in the stand. This sub-sampling yielded 3933 cases of the total 22870, or approximately 17% of out total data set.

When compared to the full dataset, it appears that the age distribution of hemlock-only dominant stands among group types is similar in age-distribution to that of the full data set (Figure 12). There are some differences that are of note. Hemlockdominated stand age distributions tends to be less skewed towards younger age categories. This is likely due to the fact that regenerated areas are replanted with more commercially attractive species such as fir and cedar. The spike visible on the left of Figure 11, in all sample populations may be the result of this silvicultural artifact.

The broader trends tend to remain the same in Figures 11 and 12. In the adjacent undefoliated, there is a weak bi-modal split between old and younger stands. Defoliated loci have a noticeably older age distribution than the undefoliated stands, and exhibit less kurtosis than subsequent non-loci defoliation.

Prior to any statistical analysis, this might suggest to an observer that older stands present a definite risk in terms of harbouring outbreak loci. It is also important to note that older stands do not have a monopoly on the initiation of an outbreak: a key point to consider when the final risk model is constructed.

Non-loci defoliation have stand-age distributions that are fairly widespread, however they do resemble adjacent undefoliated regions. This may pose complications if trying to specifically model risk for both groups L and D .

Undefoliated valleys have a noticeably younger age distribution (Figure 12), but a similar spread as all other populations. It is not unexpected to see this broad distribution in a natural population. Based upon these cursory observations, it appears that mean age may be a useful predictor of group type and hence risk, particularly for outbreak loci. The ability of age to predict subsequent defoliation and to distinguish this from undefoliated populations is less clear, particularly in undefoliated areas located within the same valley system as past outbreaks.

Appendix 5 contains the histograms for hemlock content, hemlock-cedar content, site-index, and crown closure. Initial observations do not suggest that there are distinctive differences between the various group types. They do suggest, however, that with the exception of site-index, our data require normalization prior to hypothesis testing. As some of the data exhibits a bi-modal split, this would suggest that non-parametric hypothesis testing is necessary to test if there are differences between our populations of interest. One manner of exploration is via Box-Plots confirmed through Kruskall-Wallis and Mann-Whitney U testing. Overall, considering the fact that we have such a high number of samples, this graphical analysis is one of our best tools, as all but the most ill suited variables will exhibit statistical significance.

The box-plot is a summary plot of our data. The box represents the interquartile range that contains 50% of the values. The whiskers extend to the highest and lowest values. Points beyond 1.5 box lengths are considered outliers and points beyond 3 box lengths are extreme values. The line through the box indicates the mean (Norusis, SPSS, 1994).

With regards to age (Figure 13), the initial observations of the histograms are confirmed in the boxplot observation. In the loci (L) populations, the mean age tends to be older than in non-outbreak forest stands in both adjacent (AU) and non-adjacent (UV) stands within the affected ICH (vk, wk, mw) biogeoclimatic sub-zones. Note however, that there are a large number of outliers within each of the two latter populations that may affect the ability of a model to effectively predict outbreak probability.

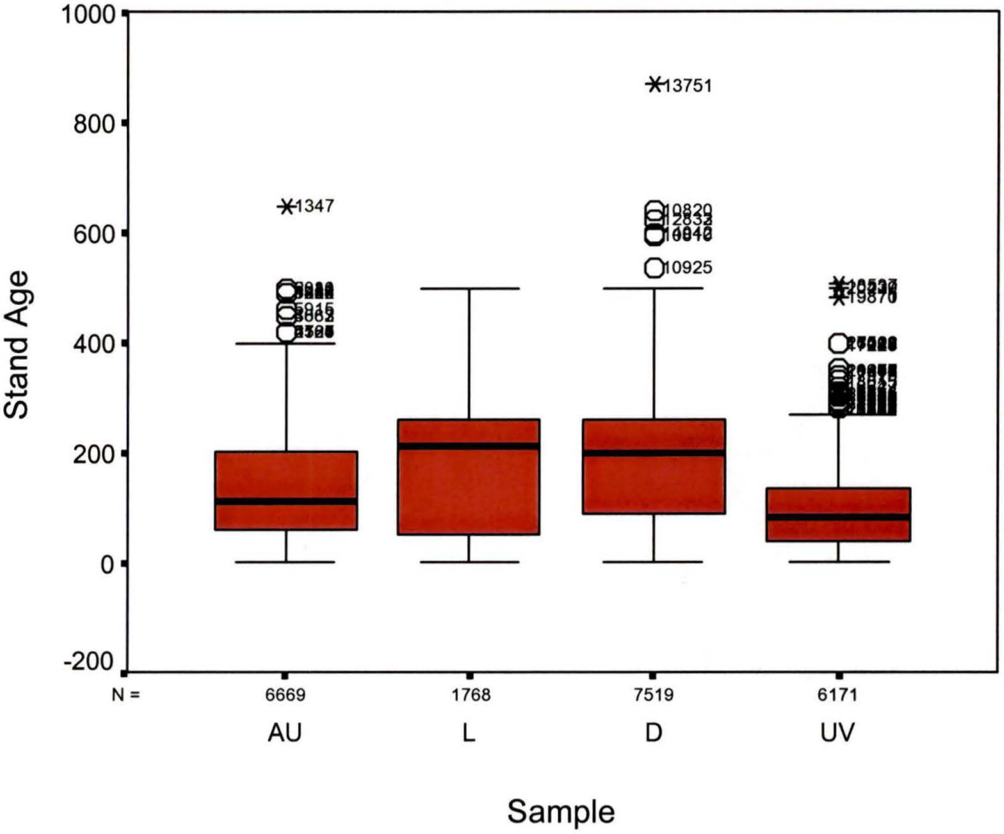


Figure 13. Box and Whisker plot of age by group type.

3.1.3 Host species

The presence of the host species is a logical feature to include in the modeling process. We must be reminded however, that since our data is derived from a forest inventory database, it is essential to test the hypothesis that the presence of a host species is both i) accurately and consistently identifiable in sufficient quantity from the database, and ii) the greater quantity of host species in a stand will increase the chances of an outbreak occurring.

Comparisons were made between the various group types with regards to hemlock content in the stand. In addition, as we are interested in a specific habitat type, western red cedar was included as an associated species of interest as hemlock and cedar are often found in close association. This may influence the micro-habitat positively for WHL compared with other associated tree species.

Figure 14 compares the means of % hemlock and % cedar content between our group types. In defoliated populations (loci and subsequent years of defoliation), both hemlock and cedar content are substantially higher than in the non-defoliated stands. Furthermore, it appears that outbreak populations of WHL occur in stands with higher % hemlock and cedar, but extend to stands with other species composition and become less discriminating as the outbreak progresses. Compared to UV, AU stands have a higher hemlock and cedar content.

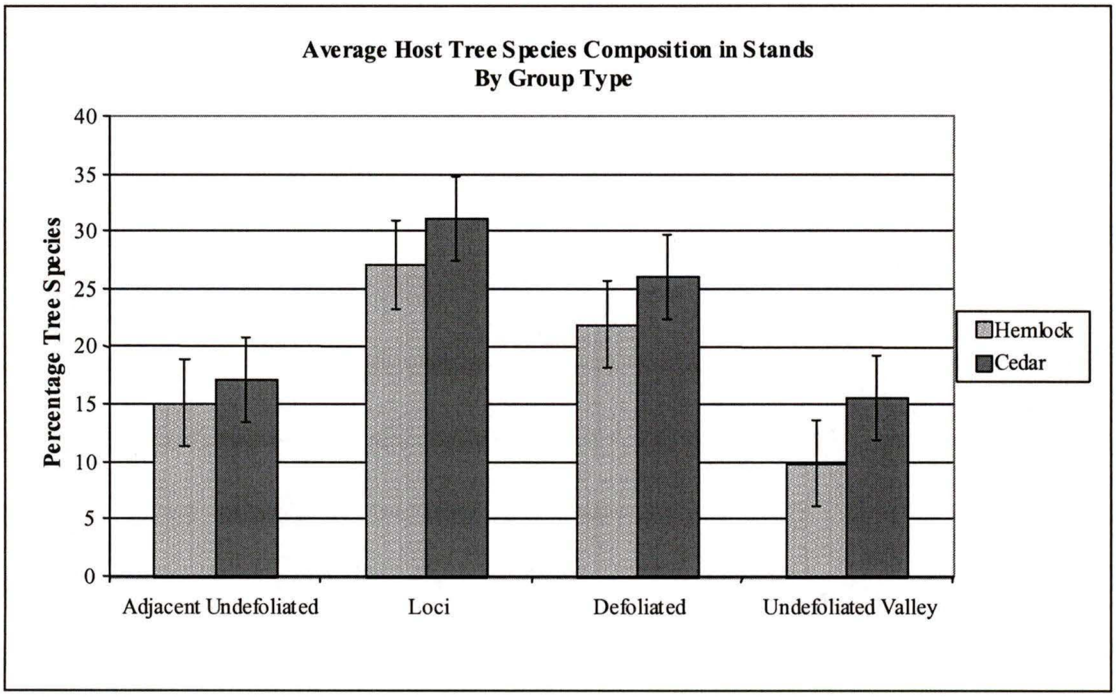


Figure 14. Mean hemlock and cedar Content

Confirmation with a Kruskall-Wallis test shows significant differences between group types (Table 8). Pair-wise multiple comparisons with a Mann-Whitney U-test also confirm significant differences (α= 0.001) between group types. These comparisons suggest that defoliation events are positively associated with the amount of hemlock content within a stand. Based upon the inventory data, the greater the amount of hemlock in a stand, the higher the chance of it harbouring the both the loci, and subsequent defoliation.

Table 8. Non-parametric tests for comparison of group types

	Mann-Whitney U-Test (Z-scores)						Kruskal-Wallis **(Chi-Sq. Dist)**
	AU vs. L	**AU vs. D**	**D vs. L**	**D vs. UV**	**UV vs. L**	**UV vs. AU**	
Age	*-12.334	-0.086	_*-21.543_	_*-48.683_	_*-28.409_	_*-26.321_	*2471.4492
Site Index	*-4.489	*-9.087	_*-16.824_	_*-52.597_	_*-30.843_	_*-31.356_	*606.3227
Air Crown	-0.4444	-0.272	-1.32	_*-23.9_	_*-10.89_	_*-17.221_	*2807.2634
Hemlock	*-21.648	*-8.433	_*-18.866_	_*-7.658_	-1.443	_*-21.459_	*461.3101
Cedar	*-25.443	*-9.733	_*-21.185_	_*-20.501_	_*-6.655_	_*-39.568_	*1471.9323
Hem-Ced.	*-48.39	*-12.229	_*-25_	_*-17.613_	_*-25.702_	_*-12.141_	*1145.1207
Avg Precip.	*-5.6909	_*-3.504_	*-11.6887	*-52.215	*-37.3538	*-46.1117	*3477.7905
Avg. Max.Temp	*-9.6803	-2.5491	*-15.8613	*-65.1704	*-45.2233	*-58.3977	*5455.0469
Avg. Min.Temp	*-11.3033	-1.7604	*-14.9406	*-64.9187	*-45.9869	*-56.3827	*5328.1797

* Significant at the $P < 0.001$ level

3.1.4 Crown closure:

Crown closure may be an important factor in determining the right habitat type for WHL outbreaks. Crown closure is only weakly negatively correlated with stand density for stands where the number of stems/ha are recorded ($r = -0.19$, $P < 0.01$, $N = 3884$).

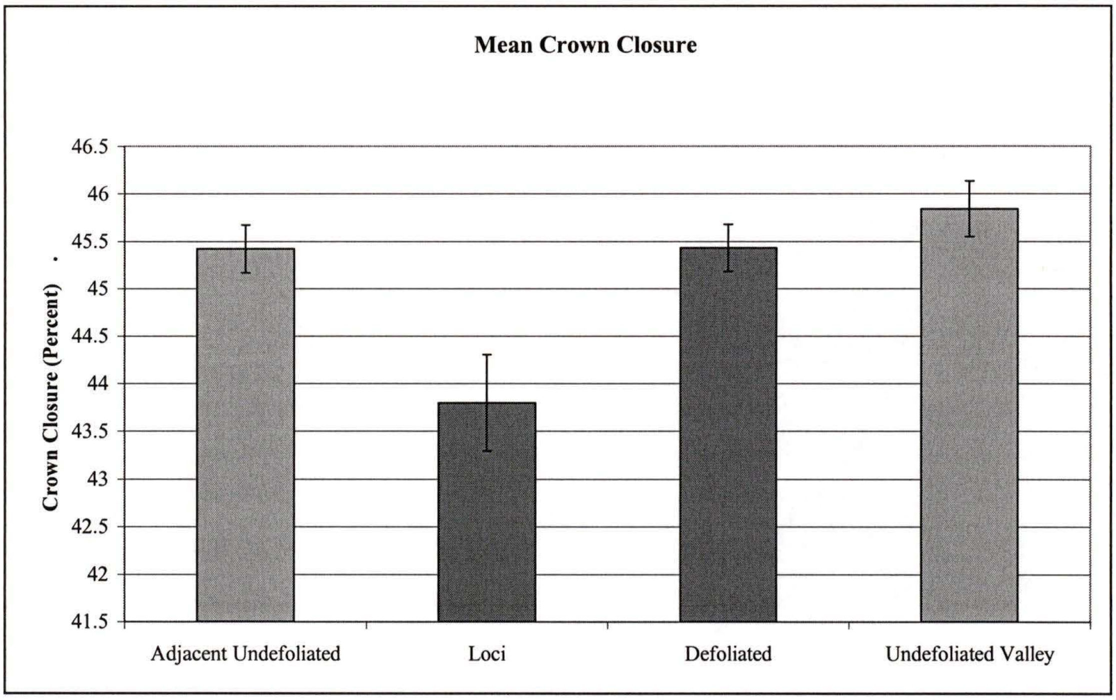


Figure 15. Mean crown closure by group type

At first instance, it appears that loci have a markedly lower mean % crown closure than the other group types. Although this might suggest that it is a useful variable in determining the locations of initial outbreaks, other factors must be considered as well. The main questions here are: i) Although it appears that there is a significant difference, U-tests confirm that there is little difference between the loci and adjacent undefoliated areas, nor subsequent defoliation, (Table 8). ii) The differences between the sample means are so small, that one must seriously question if the precision of the forest inventory would allow an individual interpreter to pick up these up from air-photos. Although statistically, one may be able to garner some predictive power from these slight differences, realistically, the original data lacks the precision to practically apply this variable and ought not to be considered in the measure of risk, as all groups fall within about 2% of one another. . This is presented as an example of the logic behind the exclusion of some variables from the analysis.

3.1.5 Site index

As it refers to site quality, one might expect that poorer sites (due to soil, water, competition, etc.) would have more stressed trees which would be more subject to defoliation events than healthy trees on healthy sites. Another hypothesis is that in areas of frequent defoliation, we would see a reduction in tree growth if the defoliation events were non-fatal. This would reflect in a lower site index. Obviously the latter situation does not present us with a causal agent, and leads us into a situation where the frequency of outbreaks, as in early hazard-rating systems, is an explanatory component of our defoliation probability. The difficulty is in teasing apart the root cause of lower site index. This raises the question of whether to include a variable if it is not in fact a causal agent.

First we must test the hypothesis that a lower site index is associated with defoliated areas. If so, this may be a useful factor in the prediction of outbreak regardless of whether it is a causal agent or simply a reaction to historical defoliation events.

From Figure 16 we see that mean site index is slightly lower overall in defoliated valleys (containing adjacent undefoliated, loci and defoliated stands), as compared to undefoliated valleys. Non-parametric hypothesis testing (Table 8), confirms this hypothesis.

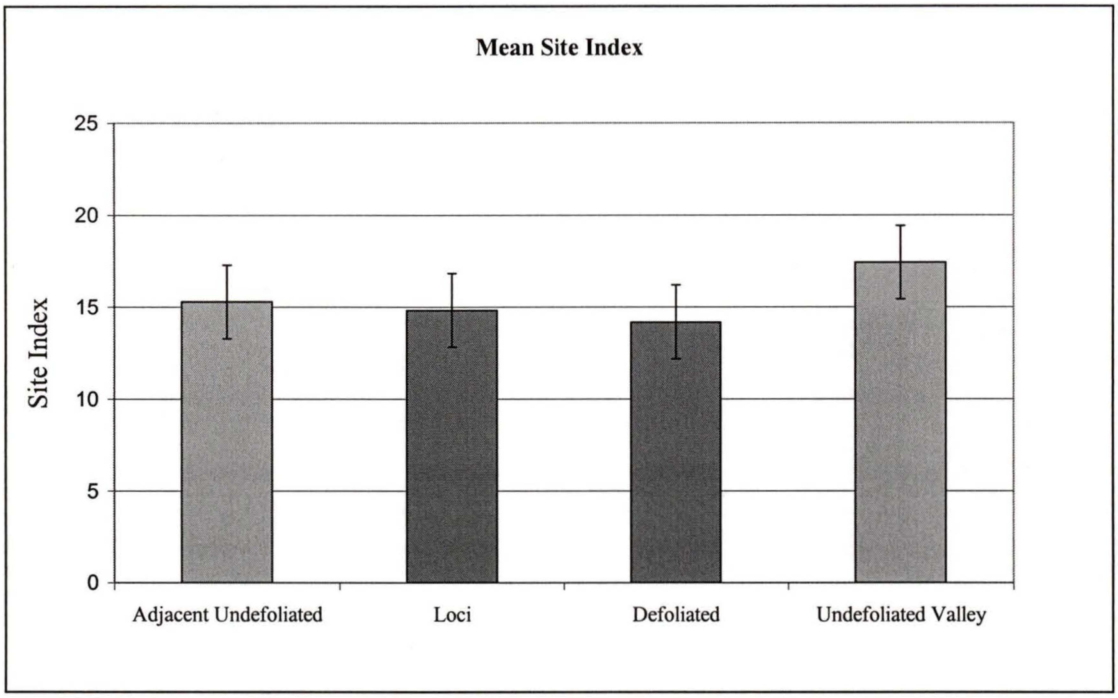


Figure 16. Mean site index

What is surprising is that loci have a higher average site index than subsequent defoliation events. This would indicate a progression to poorer quality sites as an outbreak progresses. Recall that the amount of hemlock in a stand also drops with subsequent defoliation, so that as the outbreak progresses the WHL is less discriminating for host species type and vigour of the trees it attacks.

3.1.6 Elevation, slope, and aspect

Empirical field observation of stands suggest that neither slope, nor aspect play a prominent role, as outbreaks appear to both initiate and spread along valley bottoms and part way up the hill-sides, following the presence of the favoured host species.

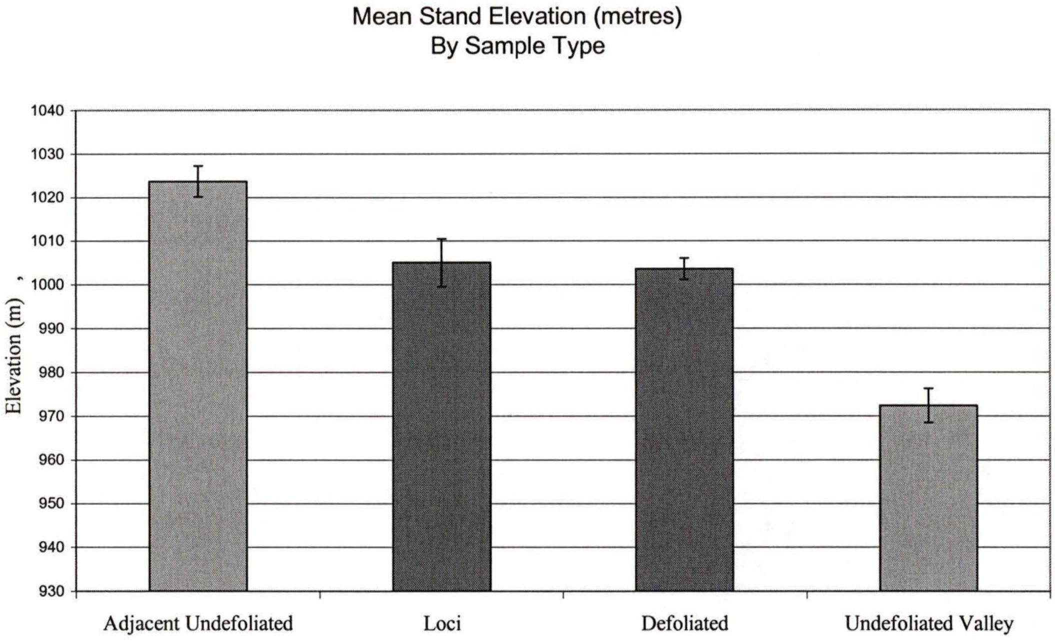


Figure 17. Average stand elevation (metres) by sample type.

On average, defoliated valleys tend to be higher in elevation by about 30 metres as compared to our undefoliated valleys. Within the defoliated valleys, as previously mentioned, defoliation occurs along the valley bottoms. This is reflected in the lower elevations of the loci and defoliated stands. This may also mean that the relative location in a valley is of importance, rather than the absolute elevation.

3.2 Results: Discriminant function analysis

The results of the discriminant function analysis are structured in the following manner:

I. Defoliated vs. undefoliated watersheds.

  A) Climatic variable model
  B) Mixed model of climate and forest cover variables
  C) Forest cover model

II. Defoliated watersheds:

  D) Discriminating among undefoliated stands, loci of the outbreak, and subsequently defoliated stands.
  E) Defoliated watersheds: discriminating between defoliated and undefoliated stands.

3.2.1 Model I: Defoliated vs. undefoliated watersheds (watershed-scale model).

Three different scenarios were run using discriminant function analysis to distinguish between defoliated and undefoliated watersheds. The first scenario explored the effects of climate variables on our ability to discern between the two watershed types. Only climate variables were entered into this first round of discriminant analysis, to examine the influence of minimum temperature, maximum temperature, and precipitation, by month. The second scenario explored the relative influence of stand information in addition to the best climate variable(s). Finally, the third run of CDF used only forest stand characteristics to distinguish between the two types of valleys.

3.2.1i A) climatic variable model

At a small scale (i.e. large area), we would expect certain characteristics such as climate to play a larger role in explaining differences between incidence of defoliation in valleys (watersheds). Overall, exploration of PRISM data clearly show that defoliated watersheds exhibit noticeably **cooler** and **wetter** conditions overall than undefoliated watersheds, despite both samples being drawn from within the same ICH-wk, -vk, and –mw biogeoclimatic subzones, Table 9.

Temperatures appear to be on average, about 2 degrees centigrade cooler within defoliated watersheds compared to undefoliated watersheds. Although there are fairly large geographic distances of 30 km or more between many of the two types of watersheds, latitudinal differences alone do not account for this fairly large divergence between the undefoliated and defoliated regions. Defoliated watersheds also appear to receive about 2 cm more precipitation on average when compared to undefoliated watersheds.

Table 9. Description of climatic variables by watershed type.

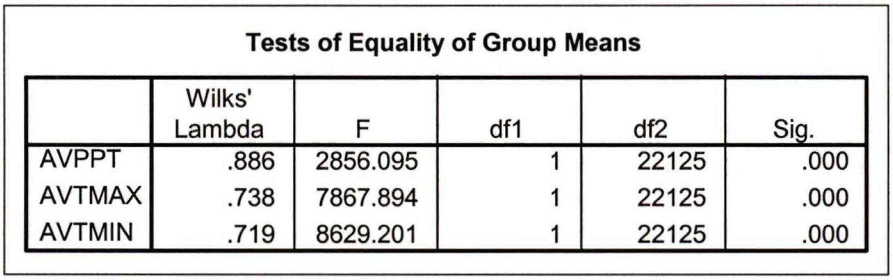
Descriptive Statistics

		N	Mean		Std.
Watershed		Statistic	Statistic	Std. Error	Statistic
Undefoliated	AVTMAX	7626	10.0950	.0192	1.67787
	AVTMIN	7626	-1.2626	.0172	1.50280
	AVPPT	7626	7.0389	.0243	2.12088
	Valid N (listwise)	7626			
Defoliated	AVTMAX	15956	8.1395	.0109	1.37955
	AVTMIN	15956	-3.0679	.0088	1.11302
	AVPPT	15956	8.9263	.0213	2.69573
	Valid N (listwise)	15956			

The first stage of running a CDF Analysis is to examine these data for normality. Most of these data were normally distributed, however log-transformations normalized those that were not. Non-parametric discriminant analysis, confirmed these results. When comparing defoliated and undefoliated watersheds using a simple t-test, there were significant differences in the mean ($p < 0.001$) in all months' temperature and precipitation with the exception of May precipitation ($p = 0.615$).

Redundancy exists within these data and as we are dealing with climate data, it is not surprising that we see a similar trend over a large area. All of our monthly climate variables exhibit a high degree of multicollinearity. To avoid this, it is necessary to choose one particular monthly variable that is representative of the other months, or to create an average of the variables. The decision was made to create 12-month averages of precipitation, minimum and maximum temperatures. This captures a signature of each valley that still is representative of the total yearly climate. This method was in agreement with models run using data from individual months.

Table 10. Watershed differences by climatic variable.



Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
AVPPT	.886	2856.095	1	22125	.000
AVTMAX	.738	7867.894	1	22125	.000
AVTMIN	.719	8629.201	1	22125	.000

Table 10 illustrates the degree to which the various climate variables differ between watershed groups. The F-statistic (as there are only two outcomes in this case, undefoliated or defoliated watershed) is the ratio of between-group variability to within-group variability. This is akin to a one-way ANOVA that compares the means of undefoliated and defoliated watersheds. Likewise, the Wilk's lambda is an index of how strongly the watersheds, UV and DV differ based upon particular variable. It is the ratio of the within-groups sum of squares to the total sum of squares. The lower the number, the stronger the contrast is between watersheds (SPSS, 1999). Table 10 illustrates that temperature plays a large role in distinguishing between watershed types, especially minimum temperatures.

Average minimum temperature has the most discriminating power between defoliated and undefoliated valleys, although it still only accounts for about 29% of the total variance in the dataset. Exploring this within the context of a CDF analysis, using average minimum temperatures yields a satisfactory model on its own without including additional variables. We see a fairly good canonical correlation (0.53) between the discriminant scores and the two watershed types; as we are only dealing with two watershed types, this is equivalent to Pearson's R (Table 11).

Table 11.Minimum temperature model summary.

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	.390[a]	100.0	100.0	.530

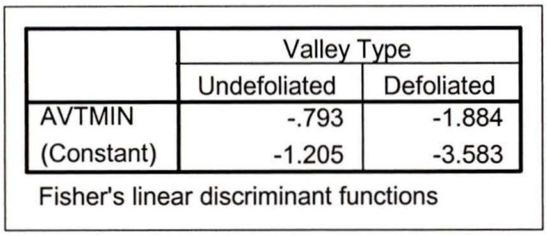
a. First 1 canonical discriminant functions were used in the analysis.

From Table 12, Fisher's linear discriminant function for this particular model is as follows:

Equation 2. Fisher's linear discriminant function for watershed scale hazard rating.

$$\mathbf{Z = -2.378 - (-1.091\ (avtmin))}$$

Table 12. Coefficients of minimum temperature model.

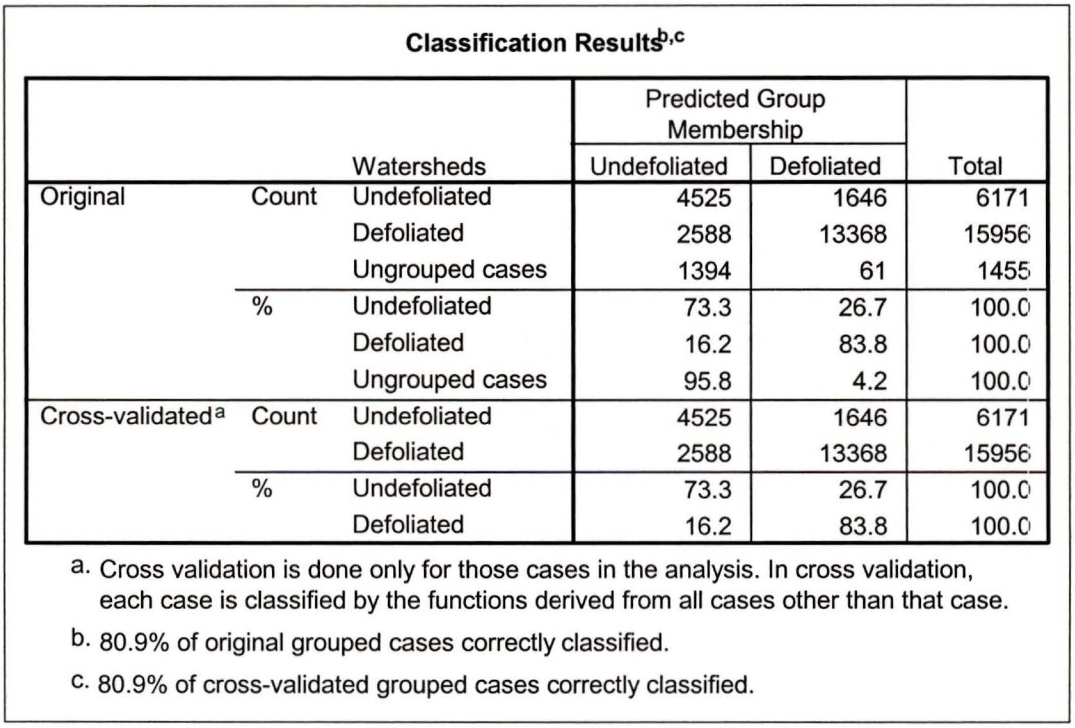


	Valley Type	
	Undefoliated	Defoliated
AVTMIN	-.793	-1.884
(Constant)	-1.205	-3.583

Fisher's linear discriminant functions

Using average minimum temperatures allows us to successfully distinguish between defoliated and undefoliated watersheds 80.9% of the time when examining the success via the leave-one-out method of model validation (Table 13).

Table 13. Model validation for minimum temperature scenario.



Classification Results[b,c]

			Predicted Group Membership		
		Watersheds	Undefoliated	Defoliated	Total
Original	Count	Undefoliated	4525	1646	6171
		Defoliated	2588	13368	15956
		Ungrouped cases	1394	61	1455
	%	Undefoliated	73.3	26.7	100.0
		Defoliated	16.2	83.8	100.0
		Ungrouped cases	95.8	4.2	100.0
Cross-validated[a]	Count	Undefoliated	4525	1646	6171
		Defoliated	2588	13368	15956
	%	Undefoliated	73.3	26.7	100.0
		Defoliated	16.2	83.8	100.0

a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

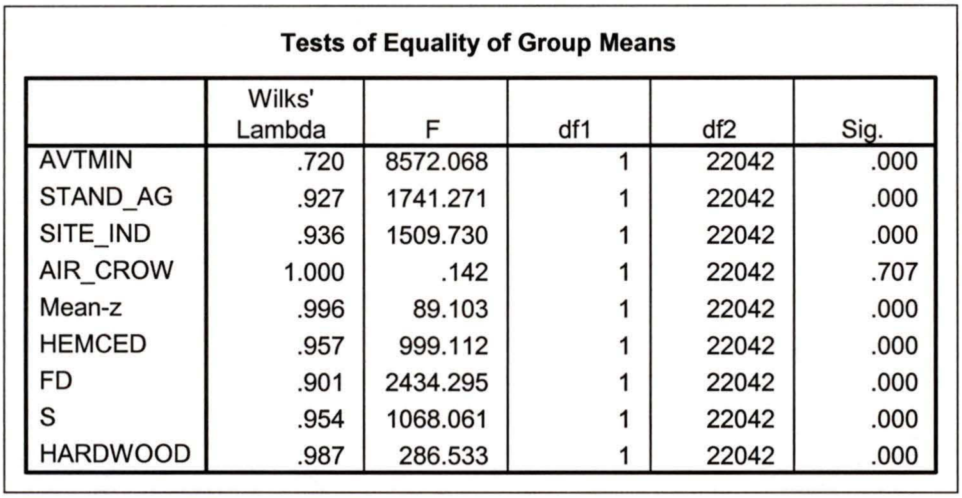
b. 80.9% of original grouped cases correctly classified.

c. 80.9% of cross-validated grouped cases correctly classified.

3.2.1ii B) Combined climate and forest cover model

In an effort to improve upon the model with climate data only, forest stand information was included. Stand variables from both undefoliated and defoliated watersheds were added. It should be noted that not all forest cover data was available for all undefoliated watersheds, and as a result, our undefoliated sample size is reduced by 1455 to 6171 stands.

Table 14. Tests of equality of watershed means

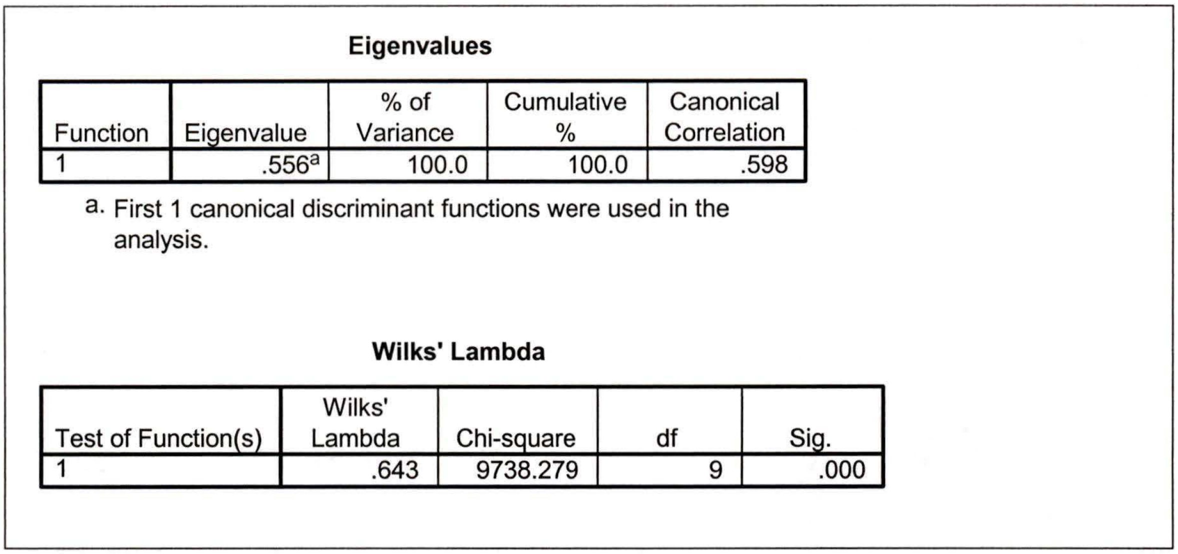


Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
AVTMIN	.720	8572.068	1	22042	.000
STAND_AG	.927	1741.271	1	22042	.000
SITE_IND	.936	1509.730	1	22042	.000
AIR_CROW	1.000	.142	1	22042	.707
Mean-z	.996	89.103	1	22042	.000
HEMCED	.957	999.112	1	22042	.000
FD	.901	2434.295	1	22042	.000
S	.954	1068.061	1	22042	.000
HARDWOOD	.987	286.533	1	22042	.000

The model improves only slightly with the addition of forest stand variables, with the canonical correlation at 0.598 (Table 15), compared to 0.530 using climatic variables alone. The new model explains approximately 36% of the variation in the data. An examination of both the standardized canonical coefficients (denoted here as β), (Table 16), and the structure matrix (Table 17) has determined the variables that have the greatest influence upon our model, as well as the direction (positive or negative) of the values that contribute to a valley that is at risk of defoliation.

Table 15. Climate-Forest cover model summary.



Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	.556[a]	100.0	100.0	.598

a. First 1 canonical discriminant functions were used in the analysis.

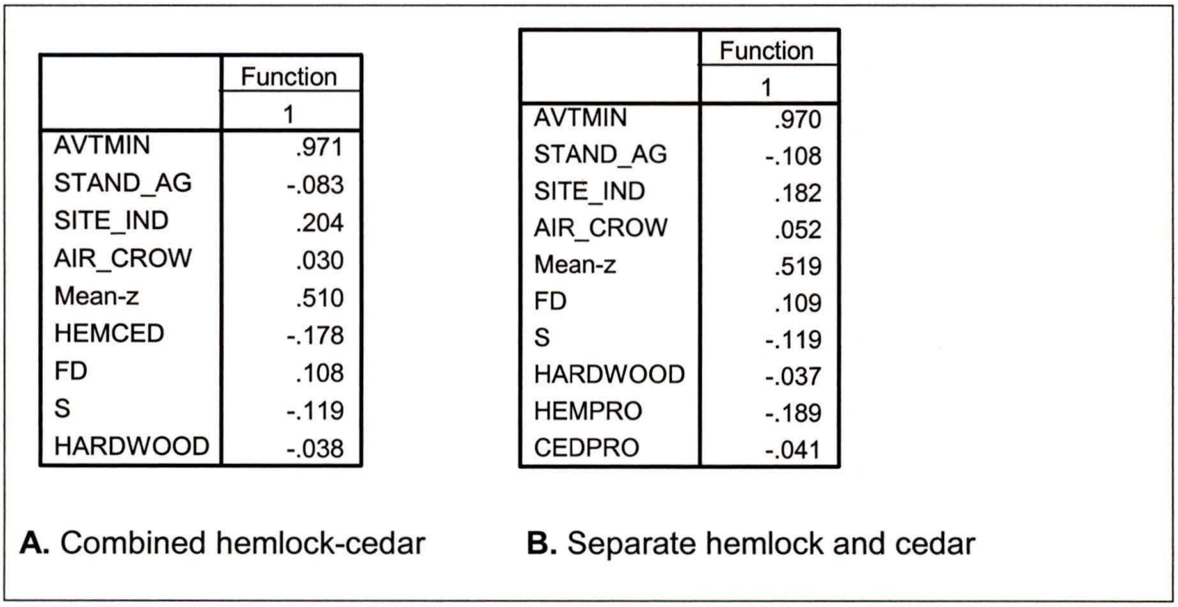
Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.643	9738.279	9	.000

Minimum temperature still has the greatest influence on the classification of our valleys, β = 0.971. The colder the minimum temperature, the greater the chance that the valley is one that is susceptible to defoliation. This is followed by elevation (mean-z), β = 0.510. Elevation is a difficult variable to interpret. Although it possesses high weighting in the overall discriminant model, it is practically uncorrelated with the resultant discriminant function. There are only slight differences in the means of the two groups, although, defoliated watersheds tend to lie somewhat higher than undefoliated watersheds. As the temperature decreases with elevation, the elevation effect may act through temperature, particularly in light of minimum temperatures carrying such high weighting in our climate model.

Site index follows elevation as the next variable of influence, β = 0.24. When we examine the structure matrix, this tends to have a clearer relationship than elevation. Lower site index stands (poorer quality sites), are more common in defoliated valleys. For stand composition, combined hemlock and cedar content play the most important role. The greater the amount of this host combination, the more likely that the stand lies within a defoliated valley. If we look at hemlock and cedar content separately, hemlock's standardized score is -0.189, and cedar is -0.041 Combined, the standardized score is .-0.178. It should be noted that both the combined hemlock-cedar and hemlock-only plus cedar-only combinations provide identical model performance. The amount of spruce within a stand has a fairly large explanatory component in our model, $\beta = -0.119$. Spruce, a species favouring moderate to wet-sites, tends to be a companion of both hemlock and cedar.

Table 16. Standardized Canonical Discriminant Function Coefficients of Climate-Forest Cover models.



	Function
	1
AVTMIN	.971
STAND_AG	-.083
SITE_IND	.204
AIR_CROW	.030
Mean-z	.510
HEMCED	-.178
FD	.108
S	-.119
HARDWOOD	-.038

A. Combined hemlock-cedar

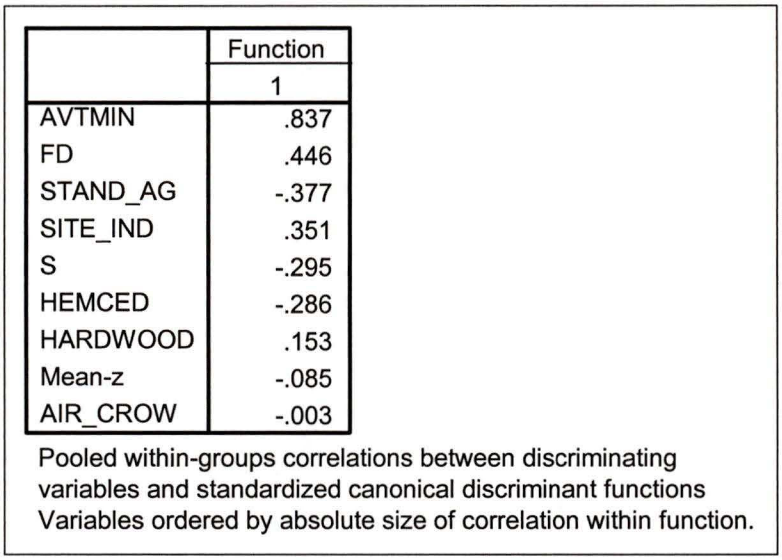
	Function
	1
AVTMIN	.970
STAND_AG	-.108
SITE_IND	.182
AIR_CROW	.052
Mean-z	.519
FD	.109
S	-.119
HARDWOOD	-.037
HEMPRO	-.189
CEDPRO	-.041

B. Separate hemlock and cedar

Douglas-fir also figures significantly and contributes to our climate-stand discriminant model, $\beta = 0.108$. The higher the proportion of Douglas-fir, (a drier-site species), within stands of a watershed, the less likely the watershed is to have experienced defoliation in the past. Hardwoods have less of a role in the model, β = -0.038, although they are negatively associated with defoliation.

Of some surprise is that stand age plays a relatively minor role, β = -0.083/ -0.108 (depending upon the model). Examination of the structure matrix (Table 17) reveals that older stands are associated with defoliated areas.

Table 17. Structure matrix of climate-forest cover model



	Function
	1
AVTMIN	.837
FD	.446
STAND_AG	-.377
SITE_IND	.351
S	-.295
HEMCED	-.286
HARDWOOD	.153
Mean-z	-.085
AIR_CROW	-.003

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions
Variables ordered by absolute size of correlation within function.

Crown closure plays an insignificant role within the discriminant model, and was excluded in the overall canonical discriminant analysis. A one-way ANOVA confirms that mean crown closure is equal between defoliated and undefoliated valleys.

Table 18. Classification function coefficients for climate-stand model

	Valley Type	
	Undefoliated	Defoliated
Mean-z	2.600E-02	2.297E-02
STAND_AG	2.068E-02	2.184E-02
SITE_IND	1.651	1.563
HEMCED	.107	.115
FD	.142	.134
S	.115	.123
HARDWOOD	.168	.171
AVTMIN	.686	-.578
(Constant)	-32.766	-31.517

Fisher's linear discriminant functions

As there are only two watershed types within this analysis, we are able to utilize the two classification functions to arrive at a linear discriminant function. This function becomes the line that separates the two groups and is calculated as:

Equation 3. Fisher's linear discriminant function for climate and stand variables at a watershed scale.

$$\mathbf{Z = -64.283 + 0.0037\ (Elevation) + 0.00496\ (Age) - 0.091\ (Site\ Index) + 0.008\ (Hemlock\text{-}Cedar) - 0.008\ (Douglas\text{-}fir) + 0.008\ (Spruce) + 0.004\ (Hardwood) - 1.264\ (Average\ Min.\ Temp)}$$

This model, while incorporating seven more variables than the minimum-temperature model, only improves our classification by another 1.3% to a total of 81.6%, using the leave-one-out method of model validation and classification (Table 19).

Table 19.Model validation for climate-forest cover model

			Predicted Group Membership		
		Watersheds	Undefoliated	Defoliated	Total
Original	Count	Undefoliated	4830	1310	6140
		Defoliated	2742	13162	15904
	%	Undefoliated	78.7	21.3	100.0
		Defoliated	17.2	82.8	100.0
Cross-validated[a]	Count	Undefoliated	4828	1312	6140
		Defoliated	2744	13160	15904
	%	Undefoliated	78.6	21.4	100.0
		Defoliated	17.3	82.7	100.0

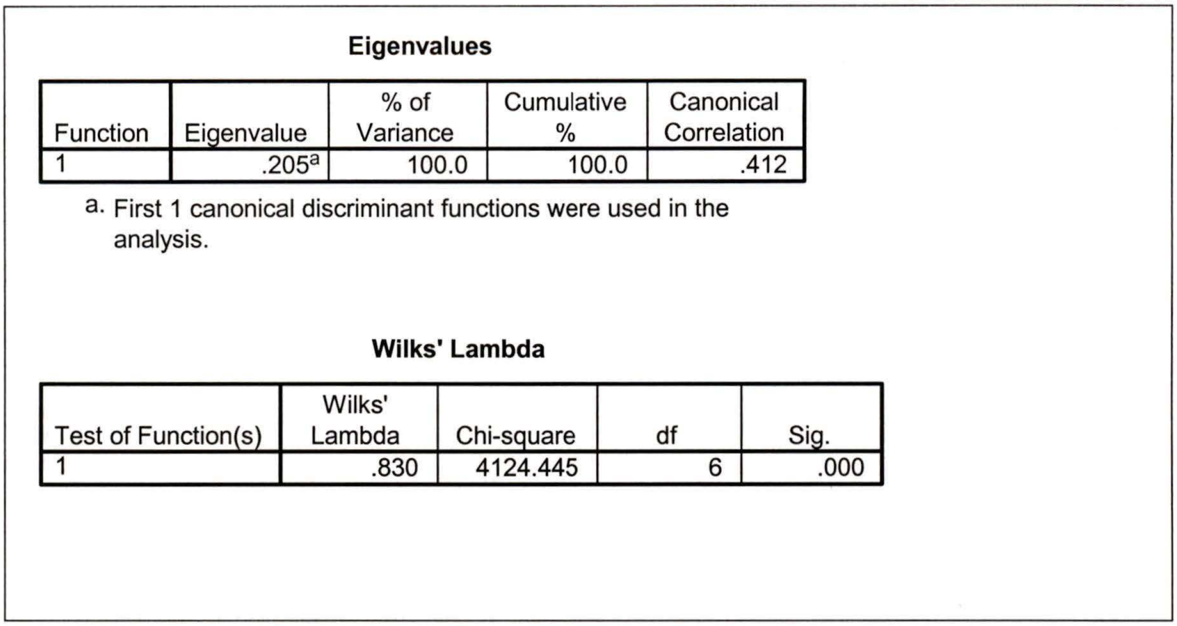
a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

3.2.1iii C) Forest cover model

It seemed prudent to examine if forest cover data alone, without the broader-scale climatic influences, could explain differences between watersheds. This could have important implications for how watershed valleys are managed. In the preceding model, forest cover data added only very slightly to our classification estimates.

When the two variables of elevation and climate are removed from the preceding model, the model's performance drops considerably. The canonical correlation is now 0.412 compared to 0.598 in the previous model, with only 17% of the total variation explained (Table 20).

Table 20. Forest cover model summary.



Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	.205[a]	100.0	100.0	.412

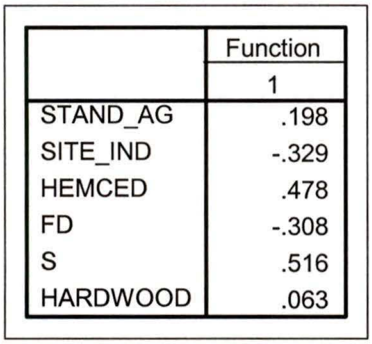
a. First 1 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.830	4124.445	6	.000

Of greater interest are the variables with the most explanatory power. The standardized beta coefficients (Table 21) show that spruce and cedar-hemlock content have the greatest effect. Site index is an important variable again, as is percent Douglas-fir. The fact that high cover of spruce and Douglas-fir appears to play strong roles in the locations of outbreaks and no outbreaks respectively, is of interest when all stands are in Interior Cedar-Hemlock biogeoclimatic sub-zones.

Table 21. Standardized canonical discriminant function coefficients of forest cover-only model.



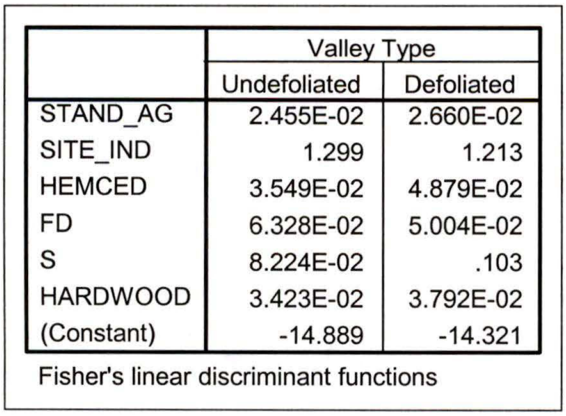
	Function
	1
STAND_AG	.198
SITE_IND	-.329
HEMCED	.478
FD	-.308
S	.516
HARDWOOD	.063

From Table 22, the resulting linear function of using stand-data alone is the following:

Equation 4. Fisher's linear discriminant function for watershed model with stand variables only.

$$\mathbf{Z = 0.02455\ (Age) - 0.086\ (Site\ Index) + 0.0133\ (Hemlock\text{-}Cedar) - 0.01324\ (Douglas\text{-}fir) + 0.06877\ (Spruce) + 0.00367\ (Hardwood) - 29.21}$$

Table 22. Classification function coefficients for climate-stand model.



	Valley Type	
	Undefoliated	Defoliated
STAND_AG	2.455E-02	2.660E-02
SITE_IND	1.299	1.213
HEMCED	3.549E-02	4.879E-02
FD	6.328E-02	5.004E-02
S	8.224E-02	.103
HARDWOOD	3.423E-02	3.792E-02
(Constant)	-14.889	-14.321

Fisher's linear discriminant functions

Overall, this model explains about 50% less of the total variation in the data than the climate-forest model. Its performance in terms of correct classification is 70.6% compared to 80.9% for average minimum temperatures alone. It is obviously a far less desirable model to predict between defoliated versus undefoliated valleys. However, it does provide us with some insight regarding the differences in forest structure that we see between the two watershed types, and provides a baseline of what one might look for when examining the differences between defoliated and undefoliated stands within defoliated watersheds.

3.2.2 Model II: Defoliated watersheds (stand scale model)

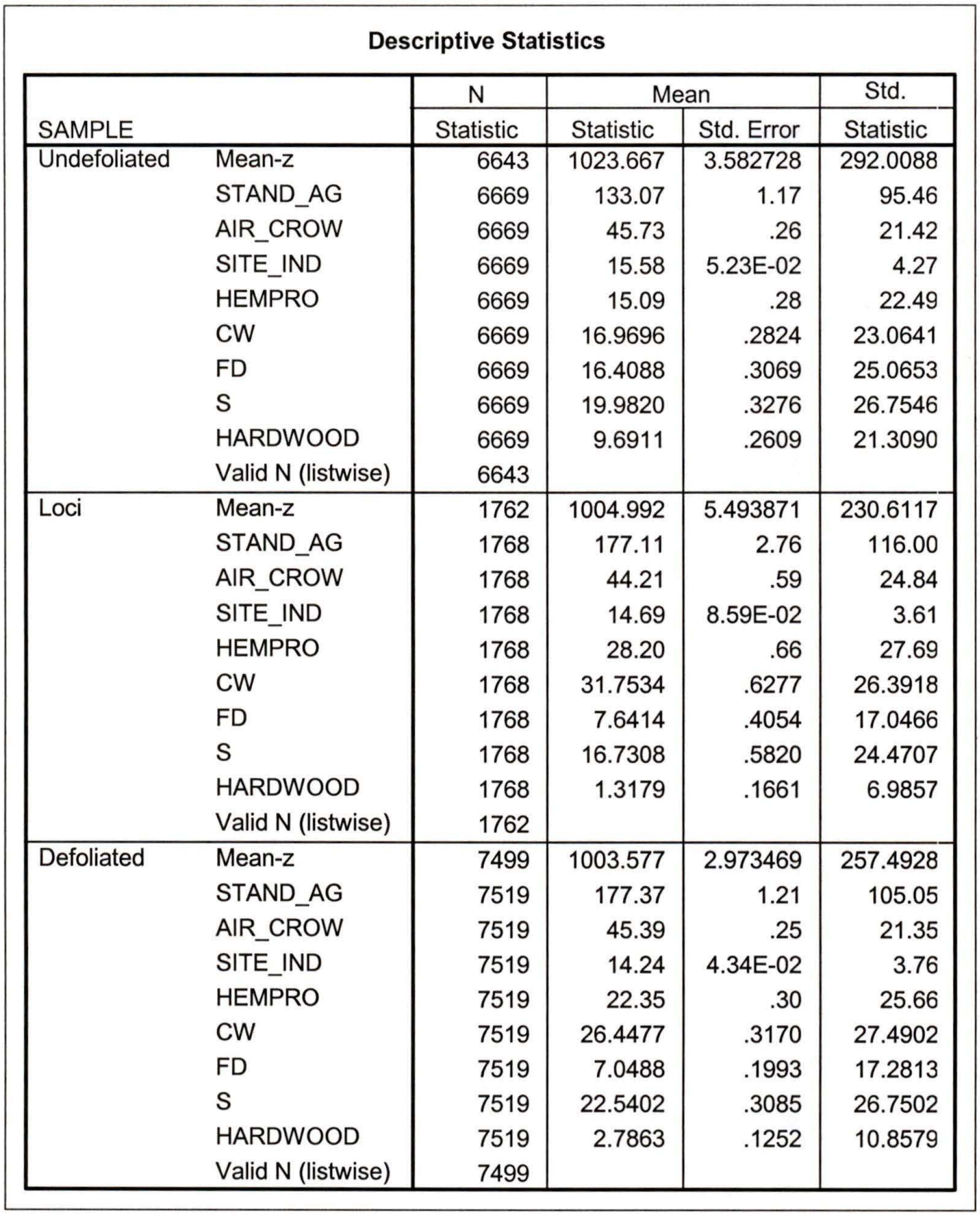
3.2.2i D) Discriminating between undefoliated stands, loci of defoliation and subsequently defoliated stands

The aim of this second series of models was to explore variables that may help to distinguish between the three group types within a defoliated watersheds and based upon the most efficient combination of variables, predict to which group unknown stands belong. Again, the aim was to find the best combination of variables that optimized the canonical correlation, Wilk's Lambda and the final results classification. Initially, 30 variables described each stand within defoliated watersheds. Variables were dropped due to multicollinearity, or because they exhibited a Wilk's Lambda close to 1; indicating no significant difference between the group means. Following the earlier modeling of stands in defoliated vs. undefoliated watersheds, the same combination of variables were explored (Table 23).

We chose to do a cursory examination of the remaining variables in our model to explore whether the observations of our variables were congruent with earlier analysis (Table 23). All values for the means appeared to match fairly well, with the exception of elevation (mean-z). Where stands in defoliated watersheds appeared to be higher than undefoliated watersheds (Model I), higher elevation no longer was associated with defoliated stands in the current model. Within defoliated watersheds, the average defoliated stand whether it be the loci of outbreak, or subsequently defoliated, was at a lower elevation than undefoliated stands by approximately 20 metres. This is not a large difference when considering the span of elevation ranges in this topologically complex area. It supports the argument that elevation does not play a large role in explaining the occurrence of defoliation between watersheds. Elevation was excluded from the analysis based upon this initial apparent contradictory relationship. CDF analysis run with this variable included also showed it to be of least importance with regards to its standardized function co-efficient. Crown closure is another variable that was dropped from the analysis, owing to the lack of difference among the groups. The result in Table 14 shows the Wilk's Lambda for Air_Crow to be 1, or not significantly different among group types.

The best model that discriminates between the three stand conditions within defoliated watersheds is very weak. The model is much weaker than CDF analysis distinguishing between watersheds. Using forest cover variables alone, we see that our canonical correlation is only 0.344 (Table 24) and explains a mere 14% of the total variation within the dataset. This would suggest that there are few differences among the three groups types, and there is little ability to distinguish among the three.

Table 23. Description of variables by group type.

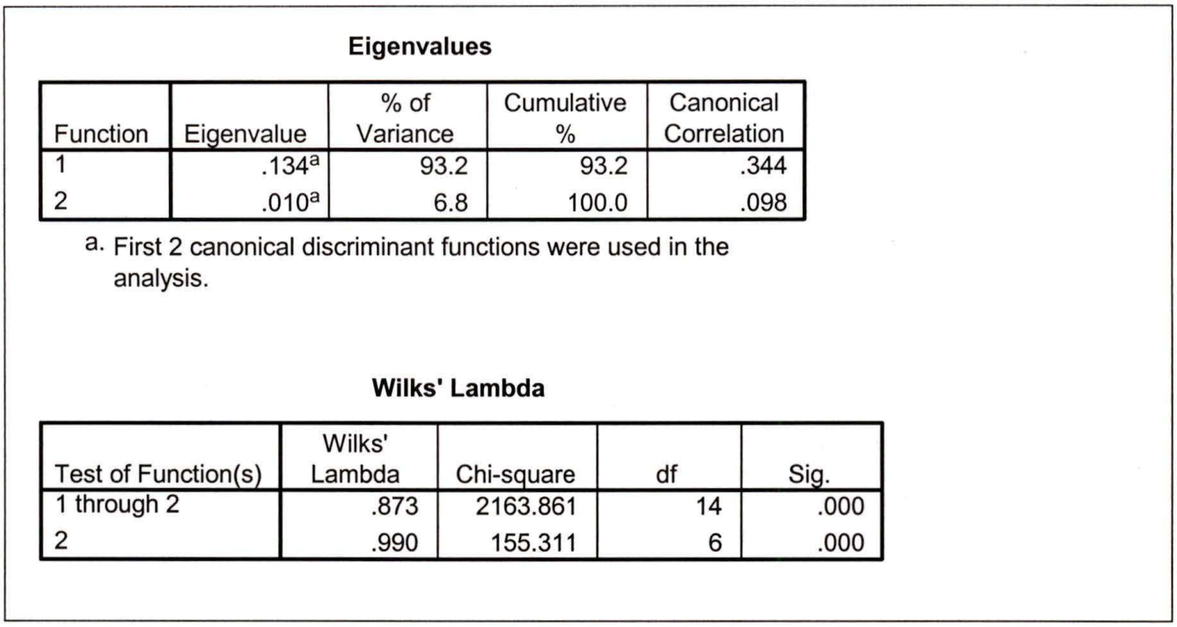


Descriptive Statistics

		N	Mean		Std.
SAMPLE		Statistic	Statistic	Std. Error	Statistic
Undefoliated	Mean-z	6643	1023.667	3.582728	292.0088
	STAND_AG	6669	133.07	1.17	95.46
	AIR_CROW	6669	45.73	.26	21.42
	SITE_IND	6669	15.58	5.23E-02	4.27
	HEMPRO	6669	15.09	.28	22.49
	CW	6669	16.9696	.2824	23.0641
	FD	6669	16.4088	.3069	25.0653
	S	6669	19.9820	.3276	26.7546
	HARDWOOD	6669	9.6911	.2609	21.3090
	Valid N (listwise)	6643			
Loci	Mean-z	1762	1004.992	5.493871	230.6117
	STAND_AG	1768	177.11	2.76	116.00
	AIR_CROW	1768	44.21	.59	24.84
	SITE_IND	1768	14.69	8.59E-02	3.61
	HEMPRO	1768	28.20	.66	27.69
	CW	1768	31.7534	.6277	26.3918
	FD	1768	7.6414	.4054	17.0466
	S	1768	16.7308	.5820	24.4707
	HARDWOOD	1768	1.3179	.1661	6.9857
	Valid N (listwise)	1762			
Defoliated	Mean-z	7499	1003.577	2.973469	257.4928
	STAND_AG	7519	177.37	1.21	105.05
	AIR_CROW	7519	45.39	.25	21.35
	SITE_IND	7519	14.24	4.34E-02	3.76
	HEMPRO	7519	22.35	.30	25.66
	CW	7519	26.4477	.3170	27.4902
	FD	7519	7.0488	.1993	17.2813
	S	7519	22.5402	.3085	26.7502
	HARDWOOD	7519	2.7863	.1252	10.8579
	Valid N (listwise)	7499			

Table 24 summarizes the three group model. The canonical correlation of 0.344 indicates that this model performs worse than the minimum temperature-only model.

Table 24. The 3-group defoliated watershed model summary



Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	.134[a]	93.2	93.2	.344
2	.010[a]	6.8	100.0	.098

a. First 2 canonical discriminant functions were used in the analysis.

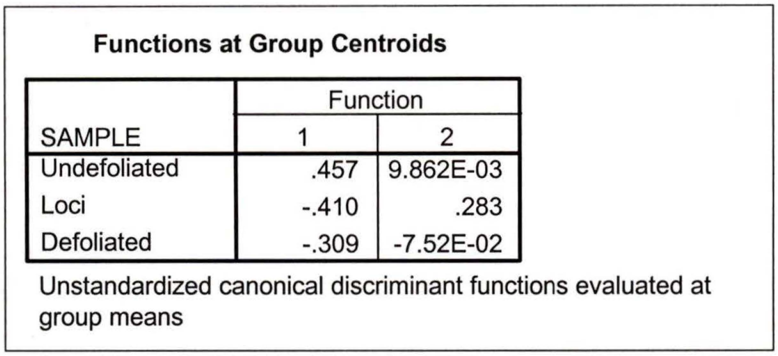
Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 2	.873	2163.861	14	.000
2	.990	155.311	6	.000

We were able to utilize the centroids (Table 25) of each group to determine how to categorize individual and unknown tree stands. The signs and values and signs of these functions also allow us to interpret the direction of the relationship for each variable.

It is worth examining the variables that do have the greatest influence on our model. The stand composition variables all weight the highest as described by the standardized CDF coefficients (Table 26). We have again utilized this in conjunction with the structure matrix to observe the direction of the relationship and the correlation of individual variables with the discriminant scores.

Table 25. Group centroids for the 3-group defoliated watershed model



Functions at Group Centroids

SAMPLE	Function 1	 2
Undefoliated	.457	9.862E-03
Loci	-.410	.283
Defoliated	-.309	-7.52E-02

Unstandardized canonical discriminant functions evaluated at group means

Cedar, hemlock and spruce are all associated with defoliation. Cedar and hemlock carry the highest weighting in our model, and have fair correlation with the discriminant scores (R = -0.57 and –0.48 respectively). The spruce has a rather low standardized score, as well as virtually no correlation with the discriminant scores. This is not surprising, considering the information presented by the means in Table 23. Spruce content is quite high in defoliated stands (Mean = 22.5%), but not within the loci (Mean = 7.6%). Undefoliated stands have, on average, 16% spruce within the stands. This information does not support a hypothesis that high spruce content leads to a greater chance of defoliation, as there is somewhat of an inconsistency among the defoliated stands within DV compared to UV (Table 16). Regionally, spruce content appears to be important, however it shows that at a local level, spruce content may only be indicative of the presence of host hemlock species on a (wet) site and the actual amount of spruce is quite insignificant. Thus, spruce is regarded as a proximal not a causal variable.

The variables that are negatively associated with defoliation are equally important to our model. Based upon our standardized beta coefficients, (Table 26), the higher the proportion of hardwood or Douglas-fir there is in a stand (indicating a drier site), the lower the likelihood that the stand will be attacked. These variables are weighted slightly higher by our model than hemlock and have the highest correlation with our discrimination function as denoted by the structure matrix (Table 27).

Site index follows stand composition variables in terms of importance. Lower site index is associated with defoliation. Stand age, interestingly enough is weighted the least, contrary to expectation. However, it is positively associated with defoliation.

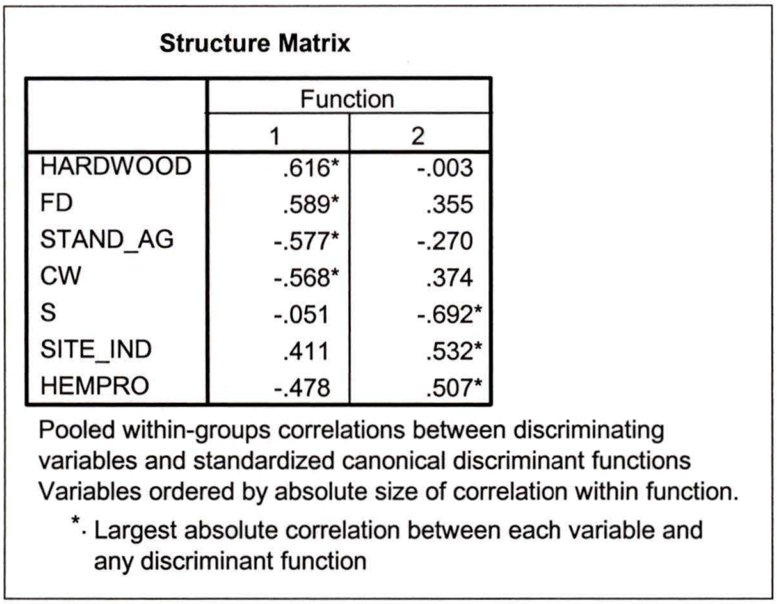
Table 26. Standardized canonical discriminant function coefficients of the 3-group defoliated watershed model.

Standardized Canonical Discriminant Function Coefficients

	Function	
	1	2
STAND_AG	-.047	-.446
SITE_IND	.233	.314
HEMPRO	-.340	.583
CW	-.452	.479
FD	.351	.284
S	-.117	-.198
HARDWOOD	.398	-.020

Both canonical discriminant functions (CDF) and Fisher's linear discriminant functions can be used to discriminate between the three group types. The first is by utilizing the canonical discriminant function centroids. We calculate the CDF function for a stand utilizing the unstandardized canonical discriminant functions (Table 28).

Table 27. Structure matrix of the 3-group defoliated watershed model



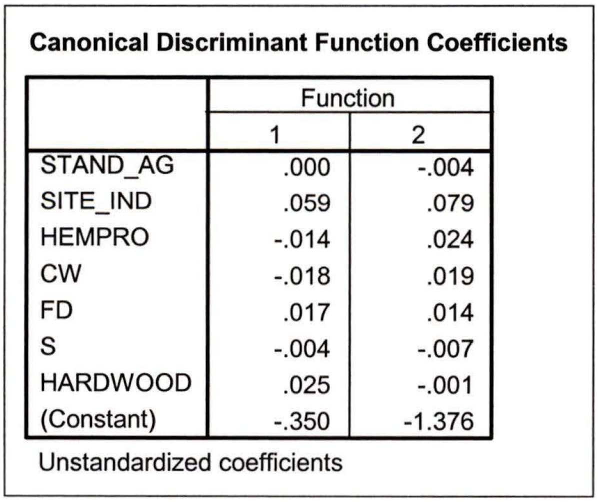
Structure Matrix

	Function	
	1	2
HARDWOOD	.616*	-.003
FD	.589*	.355
STAND_AG	-.577*	-.270
CW	-.568*	.374
S	-.051	-.692*
SITE_IND	.411	.532*
HEMPRO	-.478	.507*

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions
Variables ordered by absolute size of correlation within function.

*. Largest absolute correlation between each variable and any discriminant function

Table 28. Unstandardized canonical discriminant coefficients of the 3-group defoliated watershed model.



Canonical Discriminant Function Coefficients

	Function	
	1	2
STAND_AG	.000	-.004
SITE_IND	.059	.079
HEMPRO	-.014	.024
CW	-.018	.019
FD	.017	.014
S	-.004	-.007
HARDWOOD	.025	-.001
(Constant)	-.350	-1.376

Unstandardized coefficients

Unlike Fisher's linear discriminant functions, the canonical discriminant functions are interpreted by using the two scores (CDF computes [(n-groups) –1] discriminant functions) to place the stand in 2 dimensional space. To estimate group ownership, compare the stand's discriminant score to the group centroid. While not a true Euclidean distance, the probability that the stand belongs to a particular group is approximately proportional to what is known as the Mahalanobis distance to that group centroid. It may not be exactly proportional as we only assume a normal multivariate distribution around the centroids (Norusis, 1994). Calculation of group probabilities using Mahalanobis distances is the method that SPSS utilizes to classify cases.

Equation 5. Canonical discriminant functions for three-group stand-scale model:

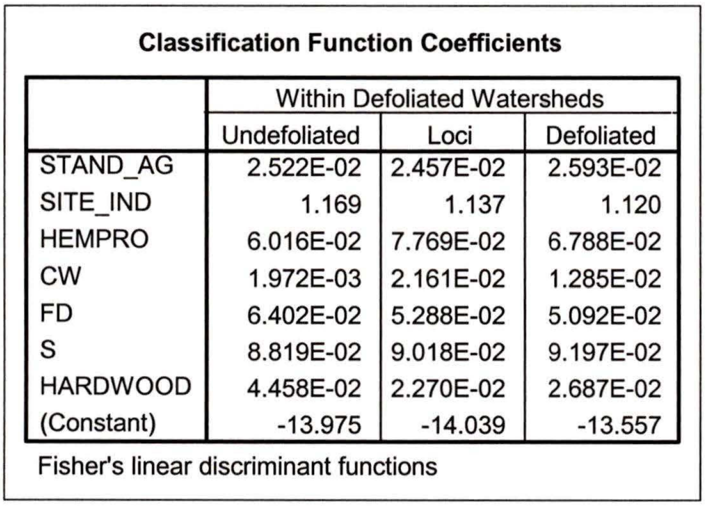
$$\mathbf{D_1 = -0.350 + 0.059(\text{Site Index}) - 0.014\,(\text{Hemlock}) - 0.018(\text{Cedar}) + 0.017\,(\text{Douglas-fir}) - 0.004\,(\text{Spruce}) + 0.025\,(\text{Hardwood})}$$

$$\mathbf{D_2 = -1.376 - 0.004(\text{Stand Age}) + 0.079\,(\text{Site Index}) + 0.024\,(\text{Hemlock}) + 0.19\,(\text{Cedar}) + 0.014\,(\text{Douglas-fir}) - 0.007\,(\text{Spruce}) - 0.001\,(\text{Hardwood})}$$

Fisher's discriminant functions are a simpler method of determining to which group an unclassified stand most likely belongs. By entering the values of the independent variables into Table 29, and examining the results for the highest value, the group membership is easily obtained, provided that the model performs well.

Examining classification success of this model is critical, and the information provided by our canonical correlation, as well as the Wilk's Lambda for the model has indicated that this model will not provide a robust of estimate.

Table 29. Fisher's classification function coefficients for the 3-group defoliated watershed model.



Classification Function Coefficients

	Within Defoliated Watersheds		
	Undefoliated	Loci	Defoliated
STAND_AG	2.522E-02	2.457E-02	2.593E-02
SITE_IND	1.169	1.137	1.120
HEMPRO	6.016E-02	7.769E-02	6.788E-02
CW	1.972E-03	2.161E-02	1.285E-02
FD	6.402E-02	5.288E-02	5.092E-02
S	8.819E-02	9.018E-02	9.197E-02
HARDWOOD	4.458E-02	2.270E-02	2.687E-02
(Constant)	-13.975	-14.039	-13.557

Fisher's linear discriminant functions

As expected from the other indicators of performance, this model has classified our stands quite poorly. Assuming equal probabilities for all cases, this discriminant model appears unable to differentiate better than by chance between defoliated and undefoliated stands. Loci are classified the best, although again, these stands are only classified correctly 58% of the time. Stands experiencing defoliation in the subsequent years of an outbreak fare the worst at a mere 32% (Table 30).

Table 30. Model validation for 3-group defoliated watershed model

Classification Results[b,c]

			Predicted Group Membership			
		SAMPLE	Undefoliated	Loci	Defoliated	Total
Original	Count	Undefoliated	3348	1747	1574	6669
		Loci	365	1041	362	1768
		Defoliated	1781	3317	2421	7519
		Ungrouped cases	4461	1188	522	6171
	%	Undefoliated	50.2	26.2	23.6	100.0
		Loci	20.6	58.9	20.5	100.0
		Defoliated	23.7	44.1	32.2	100.0
		Ungrouped cases	72.3	19.3	8.5	100.0
Cross-validated[a]	Count	Undefoliated	3342	1747	1580	6669
		Loci	366	1039	363	1768
		Defoliated	1783	3321	2415	7519
	%	Undefoliated	50.1	26.2	23.7	100.0
		Loci	20.7	58.8	20.5	100.0
		Defoliated	23.7	44.2	32.1	100.0

[a] Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

[b] 42.7% of original grouped cases correctly classified.

[c] 42.6% of cross-validated grouped cases correctly classified.

Most defoliated loci are identified in stands defoliated in years 3-5 of the outbreak, indicating that there are relatively few differences between the stands that are loci and the stands that are defoliated later. If we look at a scatterplot of the discriminant functions by group (Figure 18), we see that all groups overlap. The group centroids, indicated by the black numbers, are closest with respect to the two defoliated groups, as one might hope. All in all, stands in defoliated watersheds appear to be very similar, however one further method of grouping was explored.

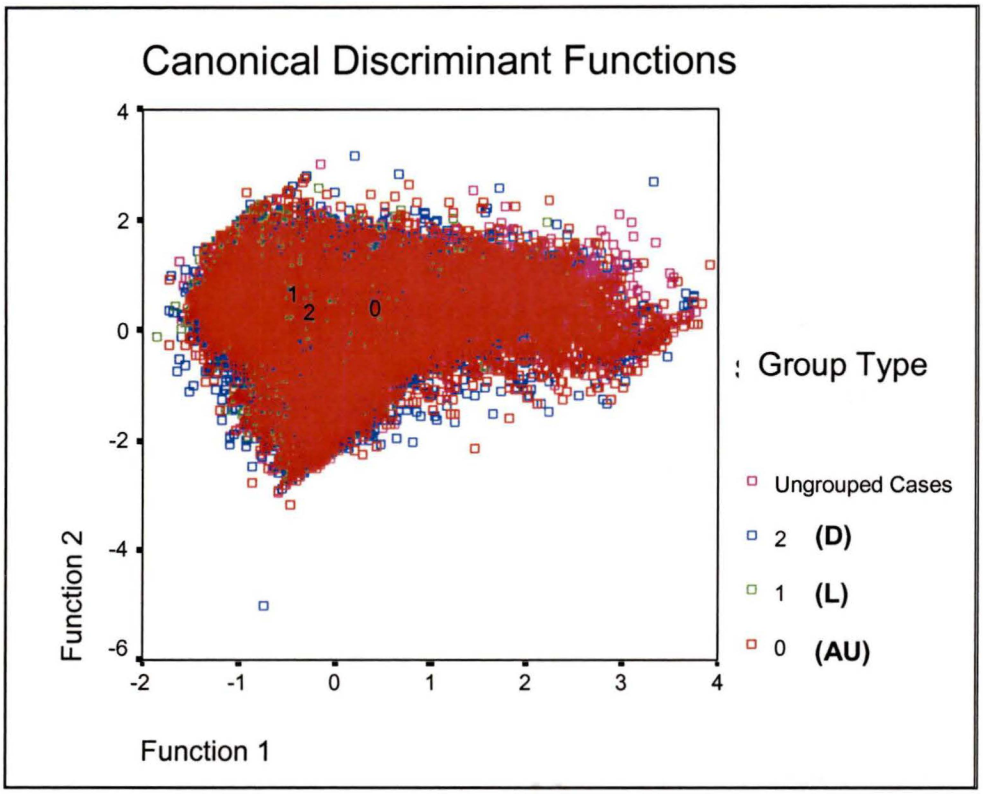


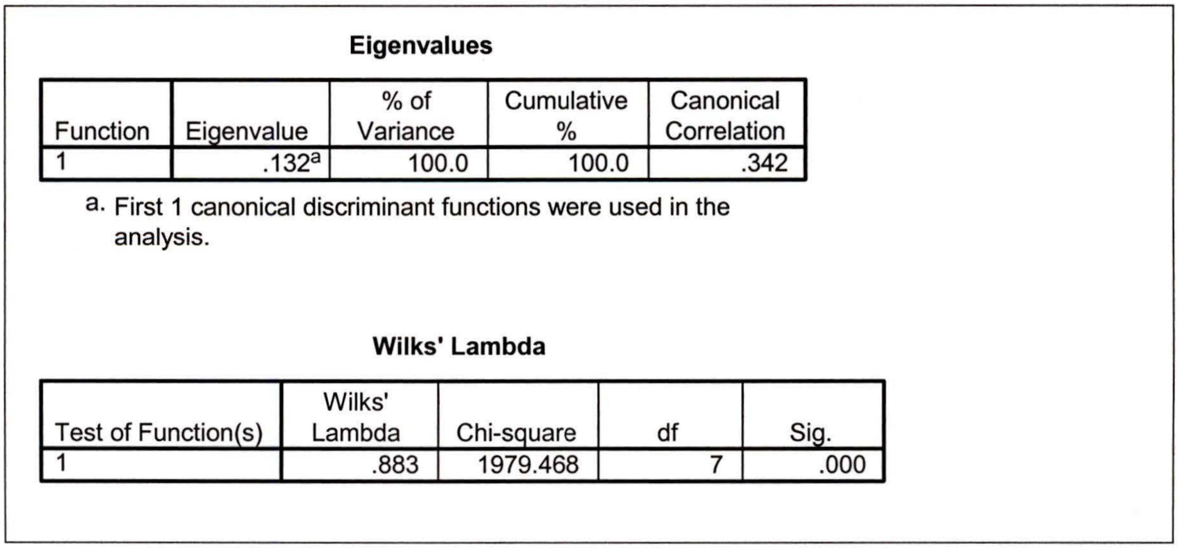
Figure 18. Scatter plot of CDA scores from the 3-group defoliated watershed model.

3.2.2ii E) Discriminating between defoliated and undefoliated stands

The previous model explored defoliated watersheds where three group types were classified based upon their stand characteristics. Loci and subsequently defoliated stands did not show a great deal of separation in either their group centroids, or the classification success of subsequently defoliated stands. It, therefore, makes sense to attempt to combine the two defoliated groups to see if it improves our model. We would expect that the fewer the groups, the lower the chance of being misclassified.

The resulting model that discriminates between defoliated and undefoliated stands within those watersheds that were defoliated also performed poorly. This two-group model only explains about 12 % of the variation in our dataset and has a canonical correlation of 0.342 (Table 31).

Table 31. The 2-group defoliated watershed model summary.



Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	.132^a	100.0	100.0	.342

a. First 1 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.883	1979.468	7	.000

It is not surprising that that we do not see a large improvement of these initial measures of model performance when moving to a two group model, considering the amount of overlap in the distributions of all the discriminant scores within the three-group model (Figure 18). The variables utilized within this model exhibit similar weighting with regards to their standardized coefficients as in the three-group model. Cedar content is weighted highest, followed by hardwood, Douglas-fir, and hemlock. Site index, spruce, and stand age wield the least influence in our model (Table 32). This is not to say that these variables are unimportant, but merely that the other variables show greater differences between the groups.

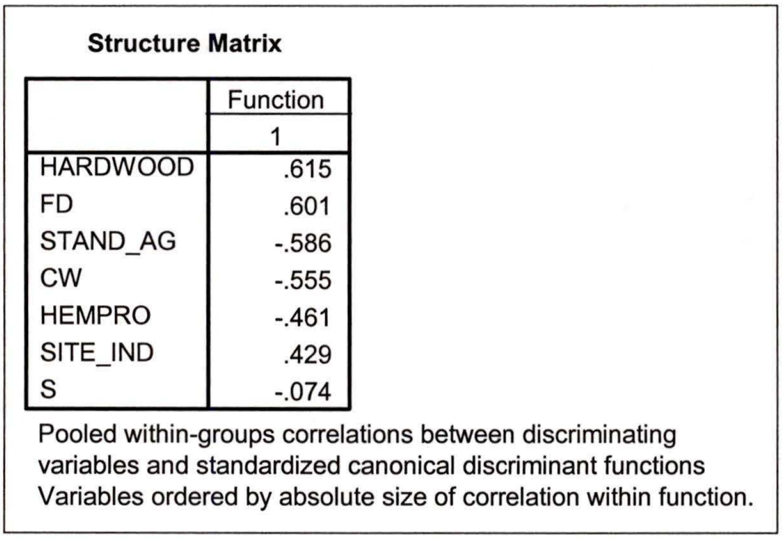
Table 32. Standardized canonical discriminant function coefficients of the 2 group defoliated watershed model.

Standardized Canonical Discriminant Function Coefficients

	Function
	1
STAND_AG	-.064
SITE_IND	.245
HEMPRO	-.318
CW	-.434
FD	.361
S	-.125
HARDWOOD	.397

Overall, our stand characteristics within the two-group model are more highly correlated with the discriminant functions than was found for the three-group model according to Table 33. Table 34 shows the means of the CDF for each group.

Table 33. Structure matrix of the 2-group defoliated watershed model.

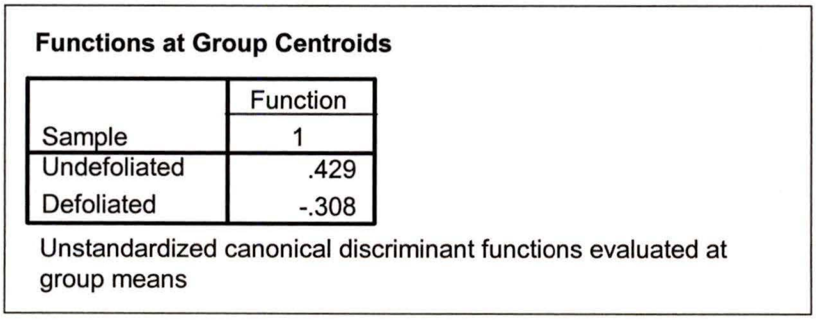


Structure Matrix

	Function
	1
HARDWOOD	.615
FD	.601
STAND_AG	-.586
CW	-.555
HEMPRO	-.461
SITE_IND	.429
S	-.074

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions
Variables ordered by absolute size of correlation within function.

Table 34. Group centroids for the 2-group defoliated watershed model.



Functions at Group Centroids

	Function
Sample	1
Undefoliated	.429
Defoliated	-.308

Unstandardized canonical discriminant functions evaluated at group means

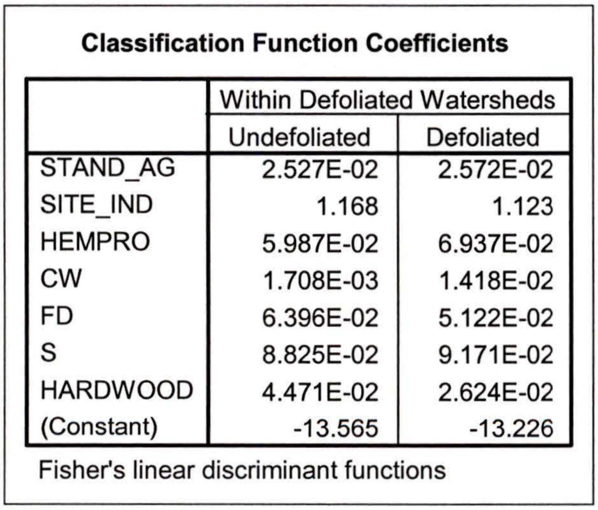
The calculated CDF model takes the following form:

Equation 6. Canonical discriminant functions for two-group stand-scale model

$$\mathbf{D_1 = -0.400 - 0.001(Stand\ Age) + 0.062\ (Site\ Index) - 0.013\ (Hemlock) - 0.017\ (Cedar) + 0.017\ (Douglas\text{-}fir) - 0.005\ (Spruce) + 0.025\ (Hardwood).}$$

Fisher's method of classification in Table 35 allows us to utilize the simplified method to classify cases. This is the recommended form to evaluate the stands in the field.

Table 35. Fisher's classification function coefficients for the 2-group defoliated watershed model.



Classification Function Coefficients

	Within Defoliated Watersheds	
	Undefoliated	Defoliated
STAND_AG	2.527E-02	2.572E-02
SITE_IND	1.168	1.123
HEMPRO	5.987E-02	6.937E-02
CW	1.708E-03	1.418E-02
FD	6.396E-02	5.122E-02
S	8.825E-02	9.171E-02
HARDWOOD	4.471E-02	2.624E-02
(Constant)	-13.565	-13.226

Fisher's linear discriminant functions

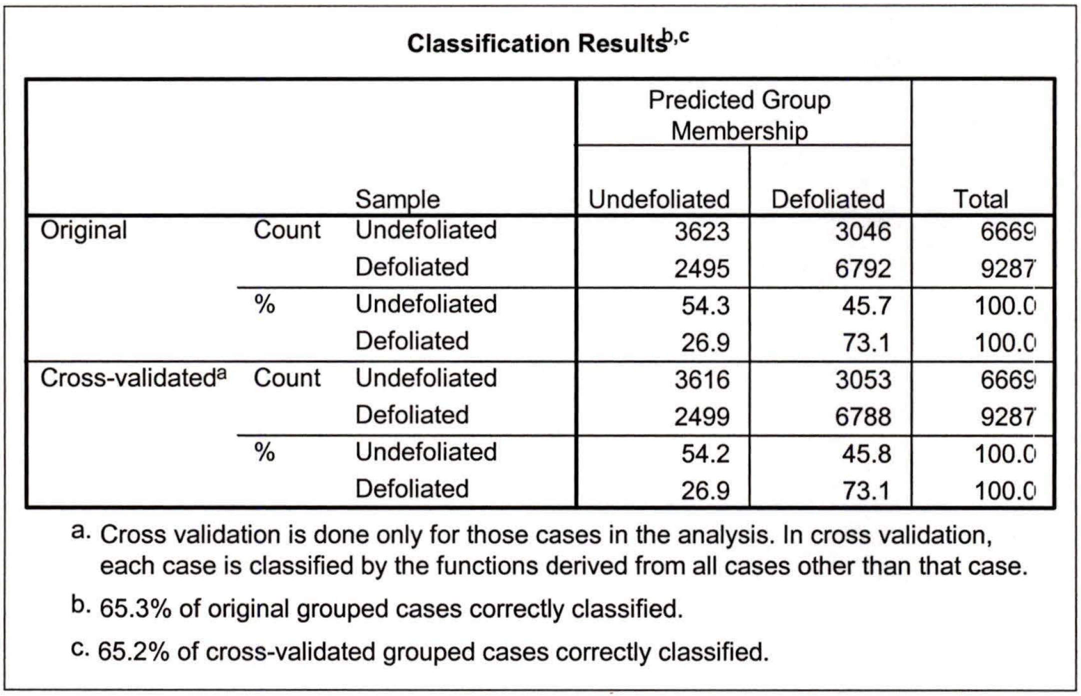
This model performs significantly better in terms of classification success, and is correct 65% of the time. The main shortcoming of this particular model is that it over-predicts stand defoliation, and misclassifies undefoliated stands about 45% of the time.

The model tends to be much better at classifying stands that have actually been defoliated, being accurate about 76% of the time (Table 36). This still is not as accurate as we would hope, however it does answer some of the questions originally posed.

Using Fisher's method of classification, we utilized a hold-back sample comprised of approximately 10% of stands not included within the original analysis, and applied the classification to this hold-back dataset for the purposes of independent model validation in addition to the standard 'leave-one out' SPSS model validation techniques. The results in Table 37 show that the model performs similarly on data that is independent of the model's construction. Again, the model drastically underestimates the number of undefoliated stands at 48%, but satisfactorily estimates the number of defoliated stands at 79% accuracy.

Overall, the two-group model is more accurate for the purposes of prediction within defoliated watersheds despite lacking the precision that we hoped to achieve with the three-group model in identifying the loci of defoliation.

Table 36. Model validation for the 2-group defoliated watershed model.



Classification Results[b,c]

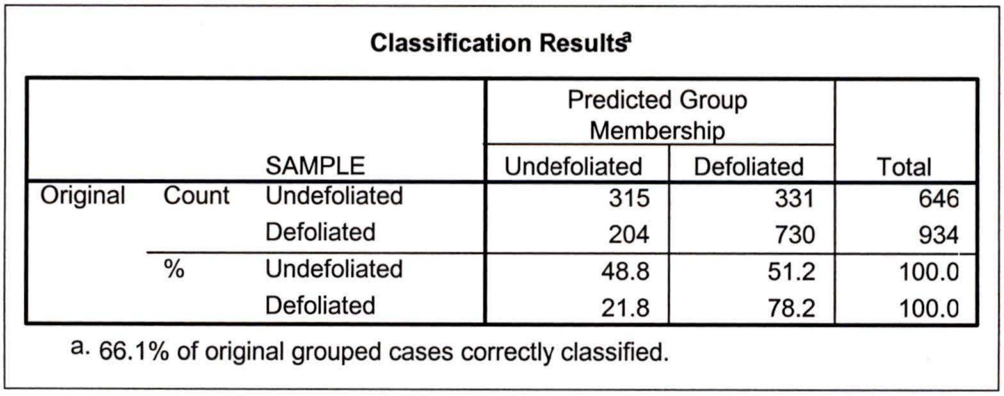
			Predicted Group Membership		
		Sample	Undefoliated	Defoliated	Total
Original	Count	Undefoliated	3623	3046	6669
		Defoliated	2495	6792	9287
	%	Undefoliated	54.3	45.7	100.0
		Defoliated	26.9	73.1	100.0
Cross-validated[a]	Count	Undefoliated	3616	3053	6669
		Defoliated	2499	6788	9287
	%	Undefoliated	54.2	45.8	100.0
		Defoliated	26.9	73.1	100.0

[a] Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

[b] 65.3% of original grouped cases correctly classified.

[c] 65.2% of cross-validated grouped cases correctly classified.

Table 37. Independent dataset model validation for the 2-group defoliated watershed model



Classification Results[a]

			Predicted Group Membership		
		SAMPLE	Undefoliated	Defoliated	Total
Original	Count	Undefoliated	315	331	646
		Defoliated	204	730	934
	%	Undefoliated	48.8	51.2	100.0
		Defoliated	21.8	78.2	100.0

[a] 66.1% of original grouped cases correctly classified.

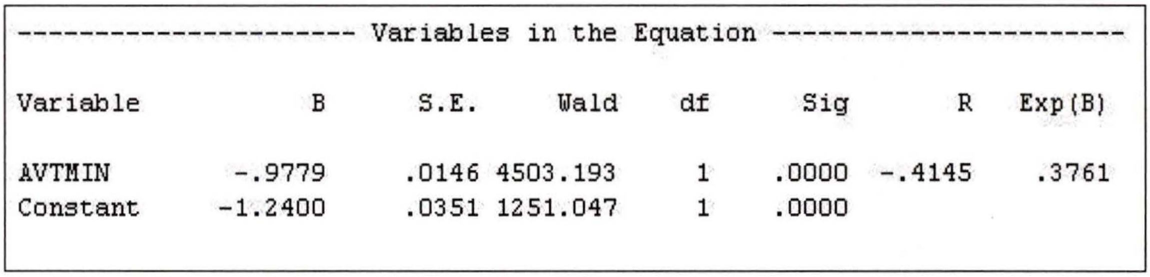
3.3 Results- Logistic regression

As there are a number of statistical methods available to the modeller, the same process performed in section 3.2 was repeated using multiple logistic regression to compare the two types of models. The modeling results were found to be in agreement. A brief summary of the modeling results is presented for the watershed and stand level models.

3.3.1 Logistic regression model of defoliated and undefoliated watersheds.

The form of this model compares stands in undefoliated valleys (UV) to stands in defoliated valleys (DV). This model results in the estimation of the probability that a stand will be defoliated within a defoliated watershed (Table 38).

Table 38. Estimation of coefficients for watershed-scale logistic model



```
---------------------- Variables in the Equation -----------------------

Variable            B        S.E.      Wald    df       Sig        R    Exp(B)

AVTMIN          -.9779      .0146  4503.193     1     .0000   -.4145     .3761
Constant       -1.2400      .0351  1251.047     1     .0000
```

This results in the model form:

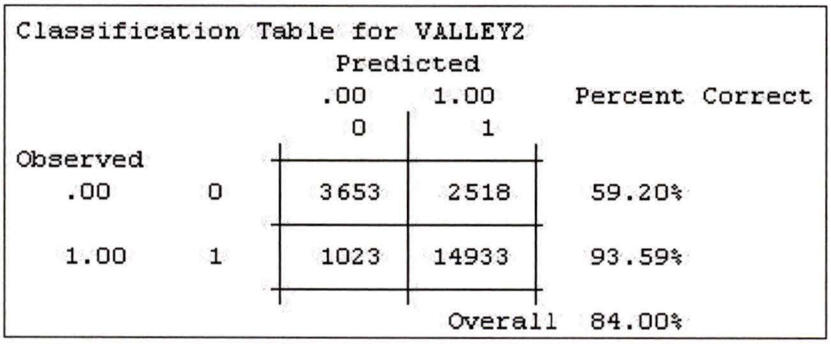
Equation 7. Watershed-scale logistic regression model.

$$\mathbf{Prob\ (defoliation)} = 1/1+e^{1.24 + 0.9779(\text{Average Minimum Temperature})}$$

The model was significant (χ^2 = 6740 df = 7, P <0.0001, -2 log-likelihood = 19 767). The Wald-statistic in Table 38 is used to test that a coefficient is 0, and has a Chi-squared distribution. The R-statistic is a partial correlation with our model and tells of a variable's relative contribution to the model. In this case, with only one variable, this statistic is not as useful as in the stand-scale model. The Exp (B) variable is an estimate of the change in the log odds, given an increment of 1 in that variable.

This particular model fares well in its overall classification of stands in defoliated valleys. Table 39 shows that this logistic model distinguishes stands in defoliated valley correctly about 94% of the time. A criticism of this model is that it misclassifies undefoliated watershed stands about 40% of the time. With an overall success rate of 84% however, this can be considered a successful predictive model.

Table 39. Classification table for watershed-scale model.

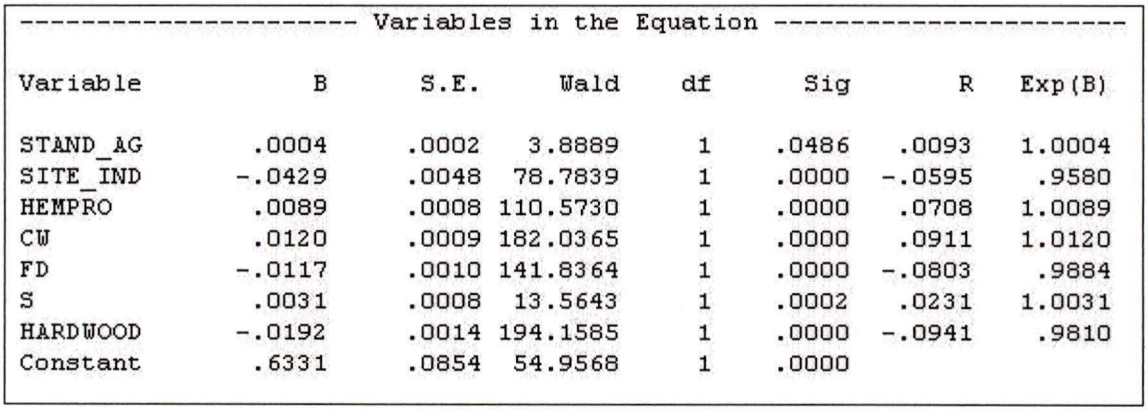


```
Classification Table for VALLEY2
                         Predicted
                        .00    1.00      Percent Correct
                         0   |   1
Observed              +------+------+
   .00         0      | 3653 | 2518 |     59.20%
                      +------+------+
    1.00       1      | 1023 |14933 |     93.59%
                      +------+------+
                             Overall     84.00%
```

3.3.2 Logistic regression model of defoliated and undefoliated stands in defoliated watersheds

The form of this model compares AU stands to combined L and D stands. This model results in the estimation of the probability that a stand will be defoliated or remain undefoliated within a defoliated watershed (Table 40).

Table 40. Estimation of coefficients for stand-scale logistic model



```
---------------------- Variables in the Equation -----------------------

Variable            B        S.E.      Wald     df       Sig        R    Exp(B)

STAND_AG        .0004       .0002    3.8889      1     .0486    .0093    1.0004
SITE_IND      -.0429       .0048   78.7839      1     .0000   -.0595     .9580
HEMPRO          .0089       .0008  110.5730      1     .0000    .0708    1.0089
CW              .0120       .0009  182.0365      1     .0000    .0911    1.0120
FD            -.0117       .0010  141.8364      1     .0000   -.0803     .9884
S               .0031       .0008   13.5643      1     .0002    .0231    1.0031
HARDWOOD      -.0192       .0014  194.1585      1     .0000   -.0941     .9810
Constant        .6331       .0854   54.9568      1     .0000
```

This results in the model form:

Equation 8. Stand-scale logistic regression model

$$\textbf{Prob (defoliation)} = 1/1+e^{-(Z)}$$

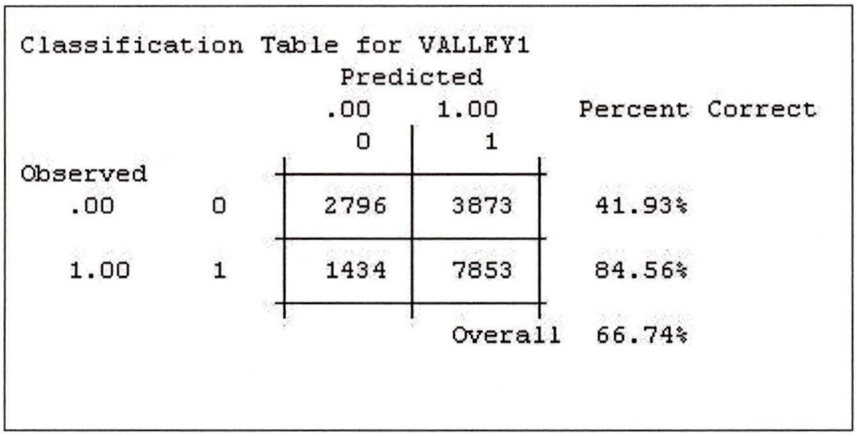
Where:

$$Z = 0.6331 + 0.0004(\text{Stand Age}) - 0.0429(\text{Site Index}) + 0.0089(\text{Hemlock}) + 0.012(\text{Cedar}) - 0.0117(\text{Douglas-fir}) + 0.0031(\text{Spruce}) - 0.0192(\text{Hardwood}).$$

The model form was significant ($\chi^2 = 1920$ $df = 1$, $P < 0.0001$, -2 log-likelihood = 19 453). Interpretation of the coefficients show all to be significant at $P < 0.05$, although stand age and spruce are weaker variables in this model. The direction of the signs and the relative contributions agree with the discriminant analysis.

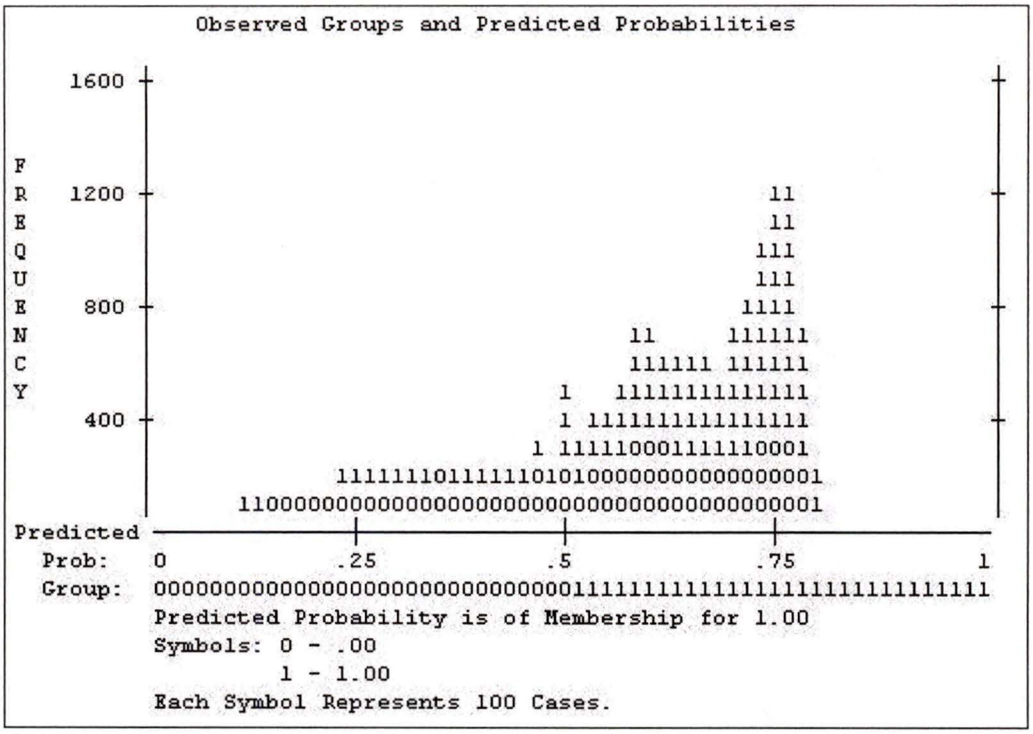
Model performance is again very good for estimating stands that are defoliated (Table 41), at 84.5 % reliability. This is at the cost of undefoliated stands however, where only 42% are predicted correctly. This is similar to the discriminant model, and indicates that there is a large degree of homogeneity between defoliated and undefoliated stands.

Table 41. Classification table for stand-scale model.



```
Classification Table for VALLEY1
                        Predicted
                      .00      1.00       Percent Correct
                       0    |    1
Observed                +---------+---------+
   .00          0       |  2796   |  3873   |    41.93%
                        +---------+---------+
   1.00         1       |  1434   |  7853   |    84.56%
                        +---------+---------+
                                  Overall       66.74%
```

This overlap in the distributions of stand qualities is apparent in the histogram of predicted probability distribution (Figure 19). Probability of defoliation is given on the x-axis; 0 = undefoliated stands, 1 = defoliated stands. Both predicted undefoliated and defoliated probabilities distributions overlap to a large degree



```
              Observed Groups and Predicted Probabilities

     1600 +                                                            +
          |                                                            |
          |                                                            |
F         |                                                            |
R    1200 +                                            11              +
E         |                                            11              |
Q         |                                           111              |
U         |                                           111              |
E     800 +                                          1111              +
N         |                                  11     111111             |
C         |                                  111111 111111             |
Y         |                             1   11111111111111             |
      400 +                             1 1111111111111111             +
          |                           1 111110001111110001             |
          |             11111110111111010100000000000000001            |
          |      110000000000000000000000000000000000000001            |
Predicted --------------+--------------+--------------+---------------
  Prob:    0            .25            .5            .75              1
  Group:   000000000000000000000000000000111111111111111111111111111111
           Predicted Probability is of Membership for 1.00
           Symbols: 0 - .00
                    1 - 1.00
           Each Symbol Represents 100 Cases.
```

Figure 19. Probability distribution for stand-scale logistic model

4 Discussion

Although this study is exploratory in nature, the discussion will present several possible biological explanations of the findings for future debate and research. The context and scope of the research is presented, followed by a summary and explanation of results, and finally, suggested directions for future research.

4.1Context

The mechanism by which widespread outbreaks of western hemlock looper occur is poorly understood. In general, only anecdotal evidence of stand and climatic attributes exists to explain the locations of large outbreaks. There is little knowledge of the associations between defoliation and the underlying forest stand structure. The goals of this study were two-fold. One was to provide a basic understanding of the characteristics of outbreaks in relation to individual stand and climate variables. The second was to build a predictive model based upon this defoliation signature of WHL.

Studies have examined stand mortality in the context of experimental plots (Alfaro *et al.* 1998). In my study however, the comprehensive GIS database of WHL outbreaks, as well as a spatially detailed, information-rich forest cover database, presented an opportunity to describe the potential outbreak signature on the landscape on a very large scale. On a smaller scale, the availability of a good BC-wide climate model made it possible to examine some of the broader trends exhibited in the location of defoliation events.

4.2 Practical application of the hazard-rating

The key to any hazard rating system is understanding how it is to be applied, as well as the relative shortcomings of the system. Foresters in the interior of British Columbia may consider evaluating the spatial susceptibility of a forest stand to WHL in a number of steps. I suggest the following process based upon the last (1990-1995) large-scale outbreak (Figure 20):

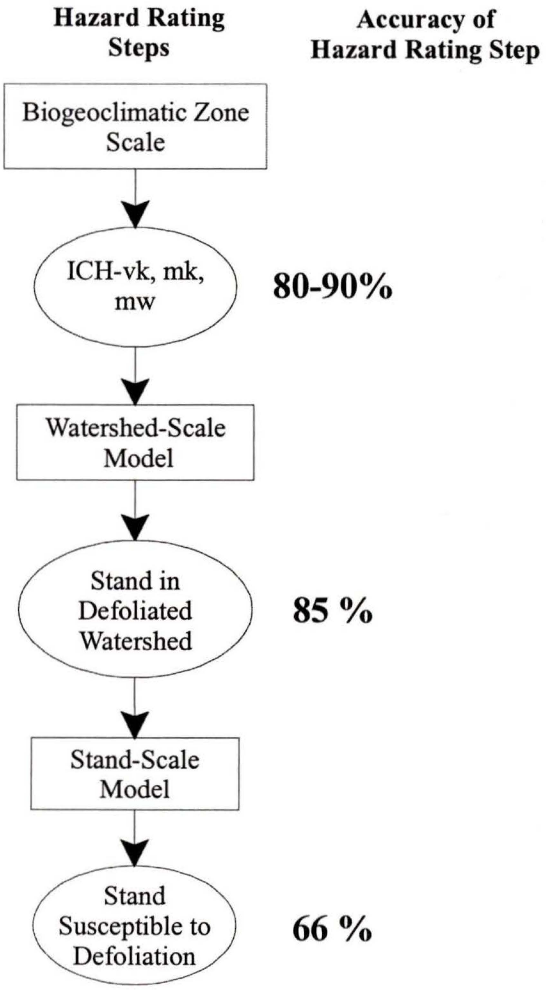


Figure 20. Hazard Rating Process

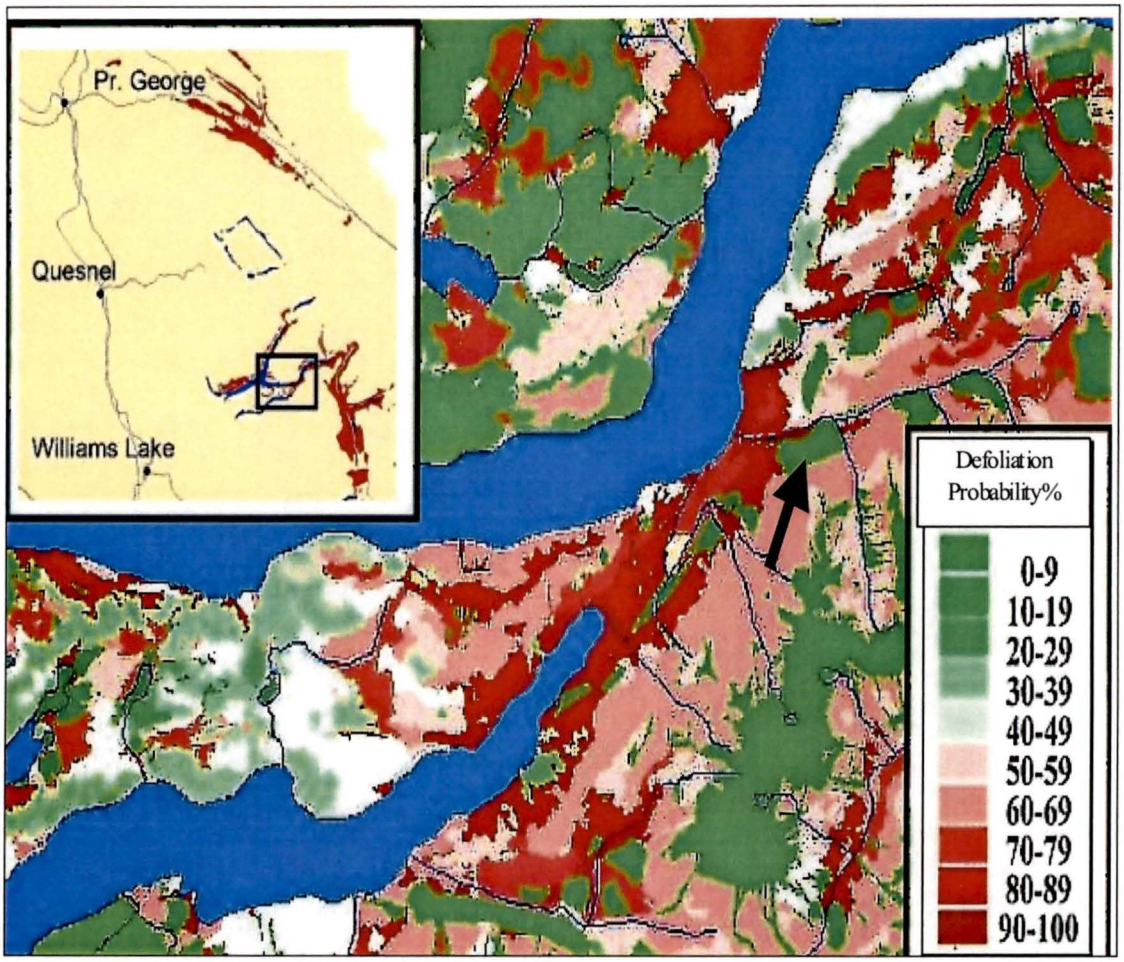


Figure 21. Stand-level susceptibility map

Should the forest manager decide to follow the entire hazard-rating process outlined in Figure 20, the stand-scale model can be applied to forest cover data and displayed in a GIS system. Figure 21 shows the stand-level defoliation model applied to a landscape in the interior of BC. This map represents an example of the end goals of this study. The red shows stands with a high probability of defoliation based upon stand characteristics, whereas the green shows a low probability. The stands most susceptible are at lower elevations in this particular map, but some low elevation stands, most notably near rivers, have a low probability of defoliation (Figure 21). The black arrow denotes one stand at low risk of defoliation because it is a young cut block. By applying these models to either the watershed scale or the stand-level scale, forest managers can gain a better understanding of areas susceptible to outbreaks and incorporate this information into the planning process prior to widespread defoliation.

4.3 Summary of main conclusions

Overall, predictions at an inter-watershed scale were far stronger than at a large intra-watershed, stand-level scale. This is somewhat disappointing, but not surprising. What it suggests is that outbreaks have some underlying small-scale controlling effects, climate being the obvious one. At a stand-level, outbreaks tend to be more stochastic in nature, owing to the homogeneity of forest stands within defoliated areas.

4.3.1 A) Watershed-scale:

i) Defoliated watersheds tend to be on average, 1.8 ° C cooler and 2 cm precipitation/month wetter than undefoliated watersheds within the same biogeoclimatic sub-zones. Average minimum temperatures provided the strongest discrimination between defoliated and undefoliated watersheds.

ii) Though minimum temperatures provided the greatest explanatory power, there were differences in forest structure (age, stand composition), between adjacent undefoliated areas (AU) and undefoliated valleys (UV).

4.3.2 B) Individual stands within defoliated watersheds

i) Detectable differences exist in the forest composition of stands defoliated in the loci and successive years of an outbreak, despite similarities in stand age. Successive defoliation is less specific to cedar-hemlock.

ii) Stand age and stand proportion of western hemlock and western red cedar are positively associated with defoliation.

iii) Stand characteristics tend to show a great deal of overlap, particularly between L and D groups. Overall, stands in defoliated valleys share very similar characteristics. Although the centroids (means) of the various samples differ within our CDA, the overlap in CDA-score distribution is very high, meaning stands tend to look the same regardless of defoliation. A clear signature is not apparent.

4.4 Suggested biological implications of results

From a practical standpoint of a forest manager, it is not essential when applying a hazard-rating model to necessarily understand the exact mechanism by which an outbreak occurs, provided that it gives accurate predictions. It is the responsibility of the researcher to attempt to provide some biological foundation for what is observed.

4.4.1Watershed-scale

Temperature is the overwhelming explanatory component when comparing defoliated vs. undefoliated valleys. Minimum temperatures alone describe 40% of the variation in the data. We can examine particular stages in the life cycle of the looper that may account for this.

4.4.1i Larval emergence

Both Ouellet (1979) and Carroll (1999) have noted the emerging first instars are heavily dependent upon fresh new foliage from the spring bud-flush. In fact, young larvae may starve if the current year's foliage is depleted (Carroll, 1956). This may be due to the fact that the mandibles of first instar larvae are physically unable to initiate feeding on older, harder needles. Carroll (1999) has observed that when provided with only older foliage, the survival rate of looper larvae is 0% for 2 year-old foliage and 5% for 1 year old foliage as compared to 55% for larvae provide with fresh foliage.

This in itself does not explain why cooler climates are favouring the looper. However, there is a possible explanation if this is combined with other evidence. Baltensweiler (1993) noted, cooler temperatures tend to synchronize larval emergence. Under artificially induced warmer microclimate conditions, larval hatching is spread out over a longer period of time rather than occurring as a short peak. In addition, warmer temperatures cause embryonic development in the egg to be prolonged, thus consuming the fixed resources available before the larvae emerge and reducing survival rates. It is more likely to observe this effect as a result of seasonal variation rather than long term climate norms. This also suggests that in warmer, undefoliated watersheds, larval emergence occurs over a longer time so that later-emerging loopers are competing with already hatched (and presumably larger) kin, to compete for a preferred resource. This is known as 'scramble competition'. In favourable years with no late freezing or extreme weather conditions, these younger/later siblings may not fare as well. This strategy does have the benefit of favouring late-hatching larvae should there be an extreme spring weather event such as freezing, high winds, or torrential rains that physically knock the early-hatchers from the trees. It may be postulated that in warmer areas with a longer period of emergence, populations may never reach critical outbreak levels, but will perhaps be more stable than in adjacent cold regions. Perhaps there may be two types of looper sub-populations with differing hatching strategies within the same watersheds.

Extrapolating from Baltensweiler's theory (1993), in the colder watersheds, total larval emergence would occur over a shorter time. This would mean that all larvae would have equal access and opportunity to initiate feeding upon the fresh foliage. If this were to occur consistently in colder areas, during years of extreme weather events, looper populations would be detrimentally affected. In those years where no harsh conditions affected loopers after larval emergence, all loopers would have equal access to a resource and thus an equal chance of survival. As Carroll (1999) pointed out, once the looper's mandibles have hardened and the first instar matures, it is easier to initiate feeding upon past year's needle growth, thus increasing the looper's chances of survival to maturity. The synchrony of bud flush and mass larval emergence may therefore be directly related to the survival of a higher percentage of larvae in cooler areas. The colder climate allows for greater synchrony in emergence, this positive feedback leads to an outbreak.

This hypothesis can be tested by comparing three-tree beating counts from undefoliated watersheds to defoliated watersheds. One would expect to see more dramatic population oscillations in colder, defoliated areas than in warmer areas. This appears to be the case, as defoliated watersheds exhibit nearly twenty times the variance that undefoliated watersheds show in larval counts from permanent sample plots ($\sigma =$ 9482.207 vs. $\sigma =$450.051). Much of this variation can be attributed to dramatic population spikes during outbreaks, however in non-outbreak years, there are noticeably fewer permanent plots that show positive looper samples within defoliated watersheds. An additional test of this hypothesis would be to compare combined hatching-feeding trials in environments with different temperature profiles to confirm the existence of this intra-specific competition within WHL.

This influence of temperature on the time course of larval emergence may also serve to explain why warmer coastal regions fail to experience the widespread outbreaks of western hemlock looper of the interior regions. Climate and weather have long been targets of investigation to explain outbreaks of various insects. Martinet (1997) reviews this climatic release argument. The patterns of WHL outbreak seem to indicate a widespread climatic phenomenon governing the locations of outbreaks. These findings may appear to contradict the assumption that warmer and drier conditions are more beneficial to the success of defoliators. In fact, Otvos has shown in the eastern sub-species of *Lambdina fiscellaria*, that outbreaks were preceded by one or two years of higher than normal temperatures and lower than normal precipitation (Otvos, 1977). These findings need not be considered contradictory. Climatic conditions and weather conditions are two distinct albeit related phenomena. Climate may be thought of as a fairly static, spatial representation of the landscape, whereas weather changes from day to day and is a temporal variable. The focus here is primarily on the location and not the timing of outbreaks. It is possible that warmer and drier weather in these cooler and wetter climates allow the looper populations to increase.

According to the literature, numerous factors suggest that cooler, wetter conditions should not favour looper populations. Heavy rains tend to disrupt moth-flight and the ability to successfully mate (Thompson, 1956). Higher humidity tends to favour fungal pathogens, albeit in combination with warmer temperatures (McDonald and Nolan, 1995). Larval development and hatching success are also hindered by cold and wet weather. These conclusions are based on occasional yearly departures from the normals for these cool, wet areas favoured by the looper and their host trees.

4.4.1ii Overwintering success

Damper conditions may work in the WHL's favour. Egg mortality may be lower in regions with more snow. Snow is an excellent insulator, and could act as a temperature buffer to for eggs buried beneath the snow within the crown's branches. This would shield eggs from extreme cold snaps that could possibly damage the eggs. Cold-dry areas might be more susceptible to winter-kill episodes than regions with more snow.

4.4.1iii Parasitioids

Parasitoid populations are also governed and affected by climate and weather. One would expect that the parasitoids have co-evolved to match their hosts' life histories. However, there is the possibility that cool wet regions offer some protection from parasitoids by putting them at a stronger disadvantage than the WHL. Experiments on some parasitoids have shown that humidity is a significant factor affecting oviposition, with higher relative humidity resulting in significantly fewer progeny (Ryan, 1974). As population growth is exponential, if a slight edge exists in favour of the looper, it may allow the WHL to increase population build up ahead of parasitic population control.

4.3.1iv ***Stand Characteristics***

The third model comparing watersheds examined only stand characteristics of defoliated and undefoliated stands. This model was able to predict the outcome accurately, roughly 70% of the time. Although this model only accounted for 17% of the total variation (vs. 29% with temperature alone), it still has value when trying to explain outbreaks within what appear to be similar valleys. As the hazard rating is refined from a biogeoclimatic sub-zone level down to a watershed level, there are subtle differences in the overall stand structures between watersheds that, individually, may not amount to a great deal, but collectively do show a significant difference.

First of all, the climate analysis showed a trend toward outbreaks within the cooler and wetter locations within the ICH-vk, -wk, and –mw biogeoclimatic zones. The stand data confirm this. Taking into account the direction of the signs (Tables 16, 17 and 21) clearly illustrates that in undefoliated stands, there are greater amounts of Douglas-fir, a dry-site favouring species. Also of note is the fact that defoliated areas have more spruce (*P. engelmannii* and *P. glauca).* Spruce favour wetter sites and might be considered a cohort to *Tsuga heterophylla* within the defoliated watersheds.

Hardwoods show a similar, but much weaker trend. The primary hardwoods are balsam poplar (*Populus balasamifera)* and trembling aspen (*Populus tremuloides*).

The amount of hemlock in a stand may also be highly important. Hemlock and cedar often grow in close conjunction, favouring similar sites. Although cedar is defoliated during outbreaks, hemlock still remains the primary host. To what degree altering the balance of hemlock and cedar would affect the probability of future outbreaks occurring is still unclear and merits future investigation. What is very clear, however, is that defoliated watersheds show that hemlock and cedar occur in much greater amounts than in undefoliated watersheds. On average, stands tend to show 58% hemlock and cedar content in stands under attack for the first year (Loci), 48% in stands under attack in secondary years of outbreaks (Defoliated), 32% within the same watershed but remain undefoliated (Adjacent Undefoliated), and only 25% hemlock and cedar content within undefoliated watersheds (Undefoliated Valleys) (Table 14). Stands tend to be much older in defoliated watersheds than in undefoliated watersheds. Loci stands are on average 174 years of age, secondary defoliation occurs in stands on average 177 years of age, undefoliated stands within these same watersheds are younger but still within the over-mature age category at 139 years of age. Those watersheds that do not get defoliated at all are vastly different in age; on average containing stands only 91 years old, i.e. 50 years younger. Not only the presence but also the abundance of over-mature host-species is a key factor in the location of outbreaks. Note that this too may be related to climate. One would expect a different fire regime within drier regions. A shorter return time of large fires would not allow suitable host trees to approach the over-mature stage in the quantity that may be required to initiate a wide-scale outbreak.

4.4.2 Stand-level hazard rating

The second goal of this study was to identify stand-specific signatures that could help identify, within a single watershed, which stands would be most likely to suffer defoliation during a WHL outbreak, i.e. what stands are at risk. It is of value to foresters to be able to identify not only the stands that would be defoliated, but also to identify the loci. In the context of all previous hazard rating work upon the western hemlock looper and other forest defoliators, this is the finest scale at which one can identify defoliation risk. To identify individual trees is both impractical given the lack of tree-specific information, and illogical considering the scale on which outbreaks occur.

The initial three-group model (UA, L, D) was disappointing, given the predictive results were less than 50%. By chance alone, one might be correct 33% of the time. However, the model allowed for approximately 43% correct prediction. It was hoped that the loci could be predicted, as this would add a valuable element to the possible identification and control of outbreaks. It appears that the looper do not spread outwards from a central locus of infestation like mountain pine beetle, for example (Stark, 1982; Safranyik *et al.,* 1980; Shore and Safranyik, 1990). Rather, the looper populations seem to bloom *in situ* like gypsy moth (Liebhold and McMannus, 1991). This is supported by the simultaneous appearance of defoliation over a large area rather than in a yearly-expanding perimeter around a few scattered hotspots. The absence of a clearly separated locus signature within the forest cover data indicates the stochastic nature of defoliation events. Figure 18 shows the centroids of our discriminant function between defoliated stands and loci to be very close together. The fact that the means of the main explanatory variables are so close, and there is such an overlap in the distributions of our discriminant scores, leads to the conclusion that there is little discernable difference between loci and subsequent defoliation.

The second model lumped loci (L) and subsequent defoliation (D) together. It sought to merely separate between those susceptible stands within an at-risk watershed, and those stands that would likely not be defoliated. This type of predictive capacity is useful so that scarce resources for monitoring and control are not wasted upon stands that are not likely to be attacked by WHL. This model was successful at predicting the correct outcome about 65% of the time. It tends to overpredict for defoliation, and this is likely an artifact of a skewed sample size (6669 (42%) undefoliated to 9287 (58%) defoliated stands within defoliated watersheds).

Both cedar and hemlock content within a stand are crucial to the prediction of defoliation. Cedar and hemlock weight highly in both the two and three group models when looking at their contributions in the standardized canonical discriminant function coefficients (Tables 26 and 32). As mentioned earlier, defoliated stands, on average contain almost double the amount of hemlock and cedar as compared to undefoliated stands. Clearly the idea of 'host gravity' plays a role in whether a stand gets defoliated. Thus, we have shown that it is the abundance and not simply the presence of host species that draw defoliation events .

Both Douglas-fir and hardwoods are important factors within the defoliated watershed model and are negatively associated with defoliation. While stands within defoliated watersheds ought to share a similar climate, the presence of Douglas-fir or hardwoods within a stand signals a micro-site that is likely drier than the surrounding region. In fact, looking at the data, there are less host species (hemlock and cedar) on these sites than on sites without Douglas-fir or hardwoods ($\mu = 33\%$ vs. $\mu = 48\%$).

Although the amount of white spruce (*Picea glauca*) present differ significantly between defoliated and undefoliated stands in the defoliated watershed model are significantly different for *Picea glauca* ($P < 0.01$), there is no significant difference when all spruce species were included.

Spruce is generally not considered an important host species, even though the looper may feed upon it when there is little else available. *P. engelmannii* and *P. glauca* were combined within the general category of "Spruce", as they share similar territories and hybridize to a large degree in the overlap of their range (Taylor, 1993). This may suggest a possible silvicultural replacement for *Tsuga sp.* and *Thuja sp.,* although this may simply involve trading one set of pest problems for another.

Site index appears as a fairly important variable within the stand-level model. As previously mentioned, the ability to interpret the role of site index is a difficult task. It is difficult to link it as a causal agent in WHL outbreaks owing to the fact that past or unknown defoliation events may be responsible for a lower site index over a wide area. Overall, it appears that site index is slightly lower for defoliated sites. Adjacent, undefoliated stands display a higher mean site-index than both defoliated groups samples, and unadjacent, undefoliated sites show a substantially higher site index than all of the other samples. Although statistically significant, we might question the precision of site indices prior to commenting on such a small, albeit significant difference.

Using the BEC method of estimating site index, accuracy is moderate at best and actual site index can vary significantly from the mean indices provided by site index tables. Many environmental factors remain uniform, however , and site index is affected by a range of conditions including:

I. Microsite differences on very wet sites and dry sites affect moisture conditions

II. Variation of depth of forest floors in cooler climates affects soil temperature and nutrient availability

III. Texture differences in soils affect nutrition and/or aeration

IV. Aspect influences temperature and moisture conditions

V. Elevation affects snow duration, snow load, temperature and moisture

VI. Latitudinal range can affect growing season and temperature conditions.

(BCMOF, 1998)

Site index, unlike age, height, or volume, is a derived value with considerable variation among sites within the ICH-wk, -vk, and -mw subzone (see Appendix 3 for 1998 estimates of BEC-Site Index Variation.). Ultimately, site index is a fairly complex variable to interpret, and individuals applying the WHL hazard-rating models may wish to consider if it is appropriate for use at a particular site.

Stand age plays a role when distinguishing between defoliated and undefoliated stands in a susceptible watershed. However, when looking at the means, it is a weaker predictor than when comparisons between watersheds are made. The diminished importance of age within defoliated watershed models could occur for two reasons. The first is that most stands are within or above the age class six, or are overmature. Forest structure will not change as dramatically between 120 and 160 years as it does between 80 and 120 years. Recall that the average age of stands in undefoliated watersheds was only around 91 years. Another reason is that the age distributions between defoliated and undefoliated stands overlap so much that it clouds the relationship.

Prior to an outbreak, if one were to walk through the woods within a watershed that is considered at risk of defoliation, it would be difficult to see much of a difference between forest stands. One might observe that defoliated stands were along the valley floors. Absolute elevation appears to be a poor estimator of defoliation (Figure 17), as defoliated stands are buffered both in lower and higher elevations by undefoliated stands. This is apparent from visual inspections of the last outbreak (Borecky, personal observation). In addition, older stands are not the only ones that get defoliated. Stands as young as 5 year silviculture replants have had defoliation within them. In fact 30% of defoliated stands are under 100 years in age. This can easily occur during large outbreaks.

The actual age of the trees may not be the causal factor or a required element for a defoliation, but the structure and conditions created by an older-growth forest in close proximity acts as a nursery for building populations. When examining the locations of these younger, defoliated stands, it is apparent that they are adjacent to old-growth regions. From Figure 11 there is an obvious spike in the age histogram of very young stands (<10 years). This seems to contradict the overall trend in the findings. Spatial analysis of these stands reveals that 90% of them share adjacency to two or more old-growth stands, and 75% of these are defoliated after year one of defoliation. Again, this gives credence to the idea of a necessary host gravity (size and contiguity) to initiate an outbreak. These are not merely isolated pockets of defoliation occurring solely in young stands.

4.5 Description of final models

The final products of these analyses are two main, scale-dependent models. The first distinguishes between defoliated and undefoliated watersheds. The second model takes the results from the first model and distinguishes between defoliated and undefoliated stands. Results for the two types of models were reported for each scale. The first consisted of a canonical discriminant function, which by itself is of little practical use to a forester. In order to utilize this, the probability must be predicted by calculating the Mahalanobis distance from the particular stand's discriminant score, to the centroids of each group outcome (i.e. defoliated vs. non-defoliated groups). The probability maps of the areas of study are calculated using this method. These probability maps detail the probability of defoliation for each of the 22 000 stands utilized within the study in Arcview shapefile digital GIS format. Probabilities are given based upon the watershed-level and the stand-level evaluation of outbreaks. Logistic regression probabilities are also included.

The second type of equations reported are Fisher's linear discriminant functions. This is simply a line that divides the distribution of the two outcomes, defoliated vs. undefoliated. It is much easier to interpret and utilize in the field by foresters as it is a yes or no answer. However, it leaves little room for extra interpretation. One enters the stand's value(s) for the variables into the linear equations (there is one equation for each possible group). The higher absolute value means that the stand in question belongs within that group.

Equation 9. Watershed-Scale Hazard Rating:

Compare:

Undefoliated = -1.205 – 0.793 (Average Yearly Minimum Temperature)
Vs.
Defoliated = -3.583 – 1.884(Average Yearly Minimum Temperature)

Equation 10. Stand-Scale Hazard Rating:

Compare:

Undefoliated = -13.565 + 0.02527(Age) + 1.168(Site Index) + 0.05987(Hemlock)

+0.001708(Cedar) + 0.06396(Douglas-fir) + 0.08825(Spruce) + 0.04471(Hardwood)

Vs.

Defoliated = -13.226 + 0.02572(Age) + 1.123 (Site Index) + 0.06937(Hemlock) +

0.01418(Cedar) + 0.05122(Douglas-fir) + 0.09171 (Spruce) + 0.02624(Hardwood)

4.6 Overall model performance and sources of improvement

4.6.1 Model performance

The first watershed-level model performed satisfactorily at an accuracy of about 81% for both the original grouped cases and the cross-validated group. It still is more generous in its prediction of defoliated stands and this is perhaps due to the vast amount of defoliation within the ICH biogeoclimatic zone. While those stands that were over-predicted have net yet suffered defoliation, it may just be a matter of time until they do, particularly if their temperature regimes lie close to the distribution of disturbed stands.

One potential question regarding the first stage of modeling is why the individual stands were used, rather than the individual watersheds themselves. As the climate modeling is on a rather small scale, many stands will share the same climate, setting up a position of potential pseudo-replication. The other option would be to model the 3rd order watersheds themselves. This would involve averaging out the climate surface for the entirety of these watersheds, thereby losing some of the detail in the information. Faced with these two compromises, the former was chosen to retain as much detail as possible in the evaluation of a particular stand. The fact that stand level variables such as age, composition, etc. were being modeled at the same time as weather made the use of the stands as a sampling unit necessary. To ensure that the conclusions were not misrepresented by a possible pseudo-replication scenario, the watersheds themselves were modeled using an average of climate. Comparisons of the means of climate in the two watershed groups using Student's t-test found agreement with the choice of model presented here.

The stand-level modeling was noticeably less successful than the smaller scale watershed model. This stand model was correct 65% of the time and underpredicting (43%) for undefoliated stands. There are several reasons why this may be the case. The first is the stochastic nature of outbreaks. The stands that are actually misclassified as suffering defoliation when they have not, essentially resemble, at least from a statistical viewpoint, the stands that have been defoliated. This suggests that the stands that remained undefoliated in the last outbreak escaped defoliation by chance alone, and are likely a target for WHL in the future. A glance at the scatterplot of the discriminant functions for these data (Figure 18) show that while the group centroids are spread out, a tremendous amount of overlap of discriminant scores exists within these groups. In field observations of these stands, defoliation appears to occur and to stop at random within stands that appear remarkably similar in age, structure and composition. Adjacent areas to defoliation are not visibly different (save for defoliation), and it gives one the impression that the looper populations would have over-run these areas too, had there not been a dramatic population crash. It was expected that the loci areas might stand out in the analysis from subsequently defoliated stands but this was not the case. The widespread homogeneity of these forests combined with the apparent stochastic nature of defoliation within these regions, make much of the forests a target for defoliation within the wettest and coolest areas of ICH wk, vk and mw.

The other explanation for the very poor explanation of variation in our data, is that stochastic variation alone does not account for all of this unexplained variation and that a key variable or several key variables were missing from our model.

4.6.2 Speculation on missing sources of variation

To get to the fine detail of hazard rating that this study hoped to achieve requires an immense amount of data. Stand-level inventory is about the finest detail that exists for the landscape. It is collected for the purposes of timber inventory and harvesting as opposed to the prediction of pest populations. What is apparent from the general trends in the available data is that host species and overmature forest seem to be extremely important and quite logical in the role of promoting outbreaks of WHL. The idea of a 'host gravity' was introduced and two main elements are likely missing from the very basic inventory data that were used in this study. One is the structure of the old-growth cedar-hemlock stands, and the other is habitat connectivity or the idea of a connected body of mature cedar-hemlock

4.6.2i Old growth stand structure

Stand structure itself is likely to play a significant role in outbreaks. Both vertical and horizontal structure may influence air-borne pheromone dispersal, well as moth flight and favourable egg-laying sites. The level of detail regarding stand structure is limited in by the FIP-FC1 data, although it is of doubtless importance.

The horizontal structure is a mix of successional stages that creates a habitat mosaic of older forest, younger forest, gaps in the canopy (containing early seral vegetation), and edge. (BCMOF 2002).

Vertical stand structure describes the top to bottom structure of a forest stand. The older grown forests seemingly favoured by the WHL outbreaks are over mature stands that are typically lower density, have multi-layered canopies and are key stands earmarked for preserving biodiversity. The BCMOF indicates some of the important attributes of such stands are:

- Intercept and retain snowfall, resulting in lower accumulations of snow on the forest floor.
- Maintain moderate stand temperatures by reducing convective and radiative heat loss. These climatically-buffered areas are often selected by ungulates as winter range.

- Provide an array of branches that are a substrate for invertebrates and arboreal lichen growth. (BCMOF, 2002)

The structural diversity of these older-growth forests provide conditions that are not captured by a simple assessment of age/species inventory. In addition, microsites with specific moisture characteristics such as drier site/ different species diversity within a stand could offer some protection/resistance from defoliation. While it may not be practical or even possible to collect this sort of information over a large area, it is likely that the structure of the forest is just as important if not more important than the information gleaned from aerial photographs province-wide.

4.6.2ii Host gravity

Earlier, the idea was introduced that it is necessary to have a specific type of habitat in a large enough volume to sustain an observable population surge of WHL. This concept was termed 'host gravity'. Small pockets of isolated mature hemlock are less likely to be defoliated by WHL than large contiguous areas. This may be due to avian predation, parasitoid pressure, etc., or simply the capacity to support a large population.

This contrasts with research conducted on *Malacosoma disstria* (Roland, 2001). A key factor positively influencing populations of forest tent caterpillar in fragmented habitats is the decreased effectiveness of the nuclear polyhedrosis virus (NPV) along edge habitat where sunlight breaks down the virus (Rothman and Roland, 1998). Since the virus does better in interior portions of the forest where it is sheltered from direct sunlight, it is postulated that this acts as a harbour for the virus and thus controls populations of *Malacosoma spp.*

The population dynamics and hosts are quite different between *Lambdina spp.* and *Malacosoma spp.* In the course of this study, it was observed that large continuous areas of old growth hemlock-cedar are showing signs of defoliation within the Interior of British Columbia. The western hemlock looper tends to favour old growth hemlock-cedar forests during outbreaks. One possibility is that *Lambdina sp.* requires a minimum size of habitat to develop a population large enough to cause visible defoliation damage. This theory was tested in a preliminary exercise whereby all stands of overmature (120years+) forests containing hemlock were selected in a GIS spatial dataset. The number of polygons (stands) was counted in defoliated vs. undefoliated watersheds. Then, common borders between adjacent stands containing this overmature, hemlock were dissolved. Thus if two similar stands containing overmature hemlock were side by side, the resulting polygons would merge into one. The resulting number of polygons was counted. Table 42 shows that in an undefoliated valleys, roughly half of the original polygons remained. The closer the fragmentation ratio is to 1, the more the stands are separated by completely distinct stands that are not particularly well suited to WHL outbreaks. The closer to zero, the greater the adjacency in the sample. In fact, in defoliated stands, over 88 % of the original stands were adjacent to one another.

Table 42. Connectivity of overmature hemlock stands (120 years+).

		Undissolved	*Dissolved*	*Fragment Ratio*
Defoliated	**# of Stands**	5889	702	0.1192
	Area (Ha)	175112.5657	183437.2914	
Undefoliated	**# of Stands**	4802	2478	0.5160
	Area (Ha)	107119.3721	200885.2102	

This is a simple exercise of dissolving polygons utilizing a GIS, but it demonstrates that there is a greater connectivity of WHL habitat amongst defoliated stands than undefoliated stands. The idea of habitat connectivity is a very complex topic with wide-reaching implications that lie beyond the scope of this thesis. What this does show, however, is that there are variables beyond what is found directly within the forest cover inventory, but that can be derived from this data that may further explain the reasons for defoliation by WHL.

4.7 Directions of future research

This section of the discussion summarizes some valid research questions raised by this thesis.

4.7.1 The effects of temperature upon WHL larval emergence and synchrony with budflush

Clearly climate has a large role in determining the locations of widespread outbreaks, even within the Interior Cedar-Hemlock biogeoclimatic sub-zones at risk. It is recommended that a WHL egg hatching and feeding bioassay be set up with a minimum of four treatments: 1) an average May temperature regime for the areas where no defoliation occurred, 2) a regime where defoliation has occurred (typically 1.8 degrees Celsius colder), 3) a treatment lower than 1.8 degrees colder (perhaps –3 °C) and 4) a treatment warmer than the locations of no defoliation (perhaps +3 °C). This would test if Baltensweiler's (1993) observations of synchrony of larval emergence hold true for WHL. Upon emergence, fresh foliage should be made available in limited quantities to the first instar larvae. Should intraspecific competition occur between early and late emerging larvae, this would give credence to the idea that colder temperatures decrease competition for scarce resources by giving all first instar larvae equal opportunity to forage. This type of experiment is an amalgamation of the work accomplished by both Baltensweiler (1993) and Carroll (1999), but would help to uncover the actual mechanism by which temperature acts on outbreaks and *Lambdina* populations. It might also serve to explain why warmer coastal locations do not see the widespread outbreaks of the Interior, despite having a clear abundance of host trees available.

4.7.2 The role of over-mature hemlock stand connectivity in determining outbreaks

Both the quantity of host and the connectedness that these overmature stands have with similar stands are likely to be important factors in determining outbreaks within the interior of the Province based upon preliminary investigations. A detailed assessment of this connectivity and a comparison between defoliated and non-defoliated watersheds is required. To do this, the watersheds themselves must be utilized and classified via concrete landscape metrics. There are ample landscape expressions available to the cartographic modeller including: spatial heterogeneity, fragmentation, edge characteristics and connectivity (Trani and Giles, 1999). Each of these categories contains specific metrics that can be statistically measured and replicated. It is recommended that the cartographic modeling program (such as FRAGSTATS) be utilized to measure connectivity, contiguity, spatial integrity, and patchiness. Fragmentation has been measured using with a host of landscape metrics, however the fragmentation index of Ripple, percent forest interior and mean patch size, and number of forest patches are recommended for investigation (*Ibid,*1999).

Based upon this, a simple t-test could determine if there are differences between the defoliated and undefoliated stands. Logistic regression could be utilized to model the probability of outbreak based upon the values of a particular watershed's landscape metrics and/or it could be added as an additional variable into the model presented in this thesis. In either case, it would lend valuable insight into the mechanism by which landscape patterns may or may not influence defoliation by WHL.

4.7.3 Timing outbreaks of WHL

A complete hazard rating system has the ability not only to predict where outbreaks are likely to occur, but also when they will occur. This in itself is a complex task and one that can be approached in two ways. It was mentioned earlier that a sentinel system of 23 pheromone traps has been established throughout the interior of British Columbia, monitoring for populations of adult WHL and has been operational since 1995. As this is the current method of census survey, it makes sense to follow this dataset throughout a complete cycle of population build up, outbreak, crash, and stability. Trap catches can be correlated with the first visible signs of defoliation to forecast future outbreaks based upon census counts of adult male moths. Currently, the system has not been operating long enough to cover a complete a full cycle.

It may also be possible to try and link this census data with an examination of weather patterns to explore a possible link between the two. Currently, the weather information and monthly departures from 30 year (1960-1990) climate normals have been collected for the following interior weather stations: Dome Creek, Nakusp, Barriere, Blue River North, Malakwa, Mica Dam and Revelstoke. These all lie within 10 km of both permanent FIDS tree-beating sample plots as well as the current pheromone traps for those areas. They also share similar topography and elevation as the pheromone trap locations, making the weather in these locations indicative of the trapping sites. A link between the timing of outbreaks and the outbreak locations is important for a full hazard-rating system. It is hoped that these three main research questions can be addressed for a clearer understanding behind the mechanism and timing of WHL outbreaks.

5 Summary and conclusions

5.1 Conclusions, recommendations and impact on the British Columbia forest sector

Defoliation hazard by WHL is highest within the wettest and coolest areas of the Interior Cedar-Hemlock, vk (very wet and cool), wk (wet and cool), and mw (mild and wet) biogeoclimatic sub-zones. Roughly 90% of defoliation occurred within these biogeoclimatic zones during the last outbreak (1990-1995). Stands with very high cedar-hemlock mixed content in age classes 6 (120 years) and above are at the greatest risk of defoliation. In areas severely impacted by the 1990-1995 outbreak, particularly within the McBride area, adult male moth census currently indicates a low pest population, most likely due to the depletion of host tree stands within the area. In other locations, a moderate population rise compared to the low levels of the looper cycle, appears to be sustaining itself.

Preventative silviculture may help to reduce future risk of losses to WHL. We recommend lowering the proportion of cedar and hemlock within managed stands and promoting mixed-species stands. Harvest schedules within these areas may not be affected, as overmature stands are considered to be most at risk. Existing overmature areas should be monitored more closely as pheromone trap catch numbers rise. The impact of heavy defoliation in these areas is reduced growth, crown-die back or tree mortality. Mortality often occurs within 2 years of the start of infestation (Alfaro *et. al.*, 1999) and can continue for up to 5 years after the infestation. In the event of an outbreak, existing harvest schedules are recommended to be flexible enough to easily move to stands at high risk thereby reducing wide-scale mortality non-recoverable losses or shortfalls/spikes within the timber-supply.

From a regional management perspective, the loss of this overmature cedar-hemlock forest over time due to the impact of WHL plus logging, will likely affect biodiversity targets set within a region. Old seral stage forest retention will likely decrease over time as a result of this important natural disturbance. Depending upon the priority of these areas (for example high-visibility locations within parks), a protection scheme may be appropriate.

5.2 Summary of modeling results

The ability to detect susceptibility of forests to WHL attacks is, at the present, is scale-dependent. As one might expect, the more detailed the scale at which the entire landscape is modeled, the greater the capacity for error. Both canonical discriminant analysis (CDA) and logistic regression performed similarly. CDA is more useful to examine the cumulative effect of all variables upon the outcome and is a powerful exploratory tool. Logistic regression provides a simpler method of arriving at probability of defoliation expressed as a percentage.

Overall, predictions at an inter-watershed scale were far stronger than within watersheds, at the stand-level scale. What this suggests is that outbreaks have some underlying small-scale controlling effects, climate being the obvious one. However, at a stand-level, outbreaks tend to be more stochastic in nature, owing to the homogeneity of forest stands within defoliated areas.

Defoliated watersheds are, on average, 1.8 ° C cooler and 2 cm precipitation/month wetter than undefoliated watersheds within the same biogeoclimatic sub-zones. Average minimum temperatures provided the strongest discrimination between defoliated and undefoliated watersheds. Watershed-scale hazard rating using climatic data is an accurate predictor of susceptibility about 84-85% of the time.

Detectable differences are observed in the type of forest composition between first year of defoliation and successive years but have similarities in age structure. Successive defoliation is less specific to cedar-hemlock, however stands in groups (L) and (D) are very similar. These stands bear closer resemblance to one another with regards to age/species composition than they do to (AU). In this respect, a weak signature can be identified using canonical discriminant analysis. Stand age and stand proportion of western hemlock , and western red cedar are positively associated with defoliation by WHL. Overall, stands in defoliated valleys share very similar characteristics. Although the centroids (mean of CDA scores) for the various samples differ within our CDA, the overlap in CDA score distribution is very high, meaning stands tend to look the same regardless of defoliation state. A strong defoliation signature is not apparent. Predictions can be made using this weak signature, but should be tempered with the understanding of the limitations of accuracy at the stand-level scale.

6 References

1. Alfaro, R., I, Taylor, S., Brown, G., and Wegwitz, E. 1998, Tree mortality caused by the western hemlock looper in landscapes of central British Columbia. For.Ecol.Mgmt., 124:291-295.

2. Andreson, R.M. and May, R.M. 1980, Infection, diseases and population cycles of forest insects. Science, 210:658-661.

3. Atkinson, A.C. 1980, A note on the generalized information criterion for a choice of model. Biometrika, 67:413-418.

4. B.C.Ministry of Forests 1995, Biodiversity Guidebook.

5. B.C.Ministry of Forests 1998, Kamloops TSA, Timber Supply Review.

6. B.C.Ministry of Forests 2001, Biogeoclimatic Ecosystem Classification. http://www.for.gov.bc.ca/research/becweb/index.htm .

7. B.C.Ministry of Forests. 2002, Stand level components of biodiversity. http://www.for.gov.bc.ca/hfp/fordev/biodiversity/module03/ .

8. Bailey, T.C. and Gatrell, A.C. 1995, Interactive Spatial Data Analysis. Longman Scientific and Technical, Essex, England, pp. 1-413.

9. Baltensweiler, W. 1993, Why the larch bud-moth cycle collapsed in the subalpine larch-cembran pine forests in the year 1990 for the first time since 1850. Oecologia, 94:62-66.

10. BC Ministry of Environment Lands Parks and Spatial Vision Consulting Ltd. 1998, User's Guide to the British Columbia Watershed/Waterbody Identifier System. Revision 2.1 BC Ministry of Environment, Lands and Parks; BC Fisheries, North Vancouver 1-95

11. BCMOF 1992, Preparation and Creation of F.R.G.I.S. Files. Forest Inventory Manual, Volume 5, Resource statistics v.1.0 8

12. Bergerud, W.A. 1996, Introduction to logistic regression models. BiomBiometrics Handbook. No. 7 B.C. Min. For. Victoria BC, Work. Pap. 26/1996 Province of British Columbia, Victoria BC. 1-147

13. Borecky, Neil and Otvos, I. S. 2001, **Coarse scale hazard rating of western hemlock looper (*Lambdina fiscellaria lugubrosa*). in British Columbia.** *In:*Proceedings Integrated management and dynamics of forest defoliating insects 1999United States Department of Agriculture, Penn.6-15

14. Bowers, W. W. 1993, **Impact of eastern hemlock looper, *Lambdina fiscellaria fiscellaria* (Guen.) on balsam fir forests in Newfoundland.** 8-3-1993Combined

Proceedings:Northeastern Forest Pest Council and 25th Annual Northeastern Forest Insect Work Conference, Lantham NY17-18

15. Bradley, G.A. 1993, Parasites of forest Lepidoptera in Canada. Part I. Subfamilies Metopiinae and Pimplinae (Hymenoptera: Ichneumonidae); Part II Subfamilies Ichneumoninae and Stenopneusticae (Hymenoptera: Ichneumonidae). Environ. Can., Can. For. Serv. Ottawa ON, Publ. No. 1336, 99pp.

16. British Columbia Ministry of Forests 1998, Site index estimates by site series for coniferous tree species in British Columbia.

17. Brooks, C.N., Heaton E., Newburn D., and Mrenlender A.M. 1999, Modeling vineyard expansion in California's north coast: developing statistical models and evaluating consequences for the surrounding oak woodland landscape. ESRI 1-21

18. Campbell, D.K. 1946, Western hemlock looper(*Lambdina fiscellaria lugubrosa* Hlst.) parasitism. Dept. Agric, Sci. Serv., Div. Ent., For. Ins. Investigations Bi-mon. Prog. Rpt. 2 (6) 3-4

19. Campbell, D.K. 1946, Western hemlock looper *In:* Summary report of the Forest Insect Survey, British Columbia and Rocky Mountain National Parks. Dept. Agric, Dom. Ent. Laboratory (105 pp. unpublished) , Vernon, BC 70-77

20. Carroll, A. L 1996, **Evaluation of the predictive vapacity of the eastern hemlock looper decision support system.** *In:*Proceedings of a Workshop at the Joint Meeting of the Entomological Societies of Canada and British Columbia<Report Num>Natural Resources Canada, Canadian Forestry Service.9-14

21. Carroll, A.L. 1999, Physiological adaptation to temporal variation in conifer foliage by a caterpillar. Can.Entomol., 131:659-669.

22. Carroll, W.J. 1956, History of the hemlock looper, *Lambdina fiscellaria fiscellaria* (Guen.)(Lepidoptera: Geometridae) in Newfoundland and notes on its biology. Can.Entomol., 88:587-599.

23. Cliff, A.D. and Ord, J.K. 1973, Spatial Autocorrelation. Pion, London,

24. Cottrell, C.B. and Monts J.S. 1976, History of population fluctuations and infestations of important forest insects in the Mainland Section, Vancouver Forest District. Env. Can., Can. For. Serv., Pacific For. Res. Centre, Internal Rpt. BC-9 (42pp) Env. Can., Victoria, B.C. 3-8

25. Crosby, D. and Baker B.H. 1966, Forest insect and disease conditions during 1966. USDA, For. Serv., Div. Timber Mgmt., Alaska Region, Rpt. 7pp. Alaska 6-6

26. Cunningham, J.C. 1970, Pathogenicity tests of nuclear polyhedrosis viruses infecting the eastern hemlock looper, *Lambdina fiscellaria fiscellaria* (Lepidoptera: Geometridae). Can.Entomol., 102:1534-1539.

27. Daly, C., Neilson, R.P., and Phillips, D.L. 1994, A statistical-topographic model for mapping climatological precipitation over mountainous terrain. J.Appl.Meteor., 33:140-158.

28. Daly, C, ., Taylor, G. H., and Gibson, W. P. 1997, **The Prism approach to mapping precipitation and temperature**. *In:*Proc. 10th AMS Conf on Applied Climatology, Amer. Meterological Soc.10-20-1997Reno, NV10-12

29. Elgee, D.E. 1963, Annual survey of forest insect diseases in the Maritime Provinces, 1952-1962. Dept. For, For. Ent. Path. Branch, For. Ent. Path. Laboratory, Interim Res. Rpt. 1-24

30. Englehardt, N.T. 1953, Deterioration of looper-killed western hemlock on lower Vancouver Island. Bi-mon. Prog. Rpt. 9(6):4 Can. Dept. Agric, Sci. Serv., For. Biol. Div.

31. Englehardt, N.T. 1957, Pathological deterioration of looper-killed western hemlock on southern Vancouver Island. For.Sci., 3:125-136.

32. Erickson, R.D. 1984, The western hemlock looper. Can. For.Serv., Pacific For. Res. Centre, Pest Leaflet FPL-21, 3pp.

33. Erickson, R.D. 1987, History of important forest insects in the Kamloops Forest Region 1912-1986. Can. For. Serv., Pacific For. Centre, FIDS Pest Rpt. 87-8, 68 pp. Victoria, BC 64-66

34. Evenden, M.L., Borden, J.H., Van Sickle, A., and Gries, G. 1995, Development of a pheromone-based monitoring system for western hemlock looper (Lepidoptera: Geometridae): effect of pheromone dose, lure age, and trap type. Envrion.Entomol., 24:923-932.

35. Finney, J.R. and Bennet, G.F. 1984, *Heterorhabditis heliothidis*: a potential biocontrol agent of agricultural and forest pests in Newfoundland. J.Agric.and For.Entomology, 1:287-295.

36. Gage, S.H., Wirth, T.M., and Simmons, G.A. 1990, Predicting regional Gypsy Moth (Lymantriidae) population trends in an expanding population using pheromone trap catch and spatial analysis. Environ.Entomol., 19:370-377.

37. Gasner, D.A., Drake D.A., Arner S.L., Hershey R.R., and King S.L. 1993, Defoliation potential of gypsy moth. Research Note NE-354 USDA Forest Service, Radnor, PA

38. Gray, D., Regniere J., and Boulet B. 1998, Forecasting defoliation by spruce budworm in Quebec. Res. Note No. 7 Canadian Forest Service, Laurentian Forestry Centre., Ste. Foy, Quebec

39. Gribko, L.S., Leibhold, A.E., and Hohn, M.E. 1995, Model to predict gypsy moth (Lepidoptera: Lymantriidae) defoliation using kriging and logistic regression. Environ.Entomol., 24:529-537.

40. Gumpertz, M.L., Wu, C., and Pye, J.M. 2000, Logistic regression for southern pine beetle outbreaks with spatial and temporal autocorrelation. Forest Science, 46:95-107.

41. Harris, J.W.E., Dawson R.F., and Brown R.G. 1982, The Western Hemlock Looper in British Columbia 1911-1980. BC-X-234 Pacific Forest Research Centre, Victoria, BC

42. Hartling, L. and Carter N. 1993, Hemlock looper in New Brunswick- survey methodology and natural history studies. Rpt. of the Twenty-first Ann. For. Pest Control Forum, Nov 15-17 , 1993 Natural Resources Can., Can. For. Serv., Ottawa, Ont. 309-312

43. Heimpel, A.M. and Angus T.A. 1959, The susceptibility of certain geometridae to crystaliferous bacteria. Dept. Agric., For. Biol. Div., Bi-mon Prog. Rpt. 15 (6):2

44. Hiscock, H.I., Hudak J., and Meades J.P. 1978, Effect of saprot on pulping qualities of balsam fir killed by the eastern hemlock looper in Newfoundland. Dept. Fish. Environ., Can. For. Serv., Newfoundland For. Res. Centre, Inf. Rept. N-X-161 , Newfoundland 1-41

45. Hohn, M.E., Leibhold, A.E., and Gribko, L.S. 1993, Geostatistical model for forecasting spatial dynamics of defoliation caused by gypsy moth (Lepidoptera: *Lymantriidae*). Environ.Entomol., 22:1066-1075.

46. Hopping, G.R., Graham K., Leech H.B., and Morgan C.V.G. 1938, *Ellopia fiscellaria lugubrosa* Hulst, a study of outbreaks in British Columbia. Ann. Rpt. Vernon For. Ins. Laboratory for the year 1938, Can. Dept. Agric, Res. Branch, For. Biol. Div. (288pp.) (unpublished) , Victoria, BC 110-122

47. Hosie, R.C. 2003, Native trees of Canada. Canada Department of Environment, Canadian Forestry Service, Ottawa, ON., pp. 1-380.

48. Hosmer, D.W. and Lemeshow, S. 1989, Applied Logistic Regression. John Wiley & Sons, New York, pp. 1-307.

49. Hudak, J., Dixon P.L., and Clarke L.J. 1988, The hemlock looper in Newfoundland in 1988. Rpt. of the Sixteenth Ann. For. Pest Control Forum, Nov. 15-17, 1988 Can. For. Serv., Ottawa, ON 105-110

50. Humphreys, N. 1993, Western Hemlock Looper in the Prince George Forest Region. FIDS Pest Report 93-15, Natural Resources Canada, Can. For. Serv. Pacific and Yukon Region, Victoria 1-2

51. Humphreys, N. 1994, Western hemlock looper in the Prince George Forest Region 1994. FIDS Pest Report. 94-15 Natural Resources Canada, Can. For. Serv., Pacific and Yukon Region., Victoria 1-3

52. Jardine, A.K. 1969, Western hemlock looper in British Columbia. Dept. Fish. For., Can. For. Serv., Pacific For. Res. Laboratory, For. Pest Leaflet 21, 1-5

53. Jobin, L. and Desaulniers R. 1981, Results of aerial spraying in 1972 and 1973 to control the eastern hemlock looper *(Lambdina fiscellaria fiscellaria* (Guen.) on Anticosti Island. Environ. Can., Can. For. Serv., Laurentian For. Res. Centre, LAU-X-49E 1-30

54. Jobin, L. and Coolombe C. 1992, The luminoc insect trap. For. Can., Quebec Region, Inf. Leaflet, LFC 26 1-12

55. Johnson, N.E., Shea, K.R., and Johnsey, R.L. 1970, Mortality and deterioration of looper-killed hemlock in western Washington. J.For., 68:162-163.

56. Kimoto, T. and Carlson J. 2001, 2000 Western hemlock looper populations and potential impacts to greater vancouver regional district watershed management objectives. Report for Greater Vancouver Regional District, Policy and Planning Department Phero Tech Inc. 1-26

57. Kinghorn, J.M. 1952, Western Hemlock looper egg sampling. Dept. Agric., Sci. Serv., Div. For. Biol., Bi-mon. Prog. Rpt. 8(3) 3-4

58. Kinghorn, J.M. 1954, The influence of stand composition on the mortality of various conifers caused by defoliation by the western hemlock looper (*Ellopia fiscellaria lugubrosa*) on Vancouver Island, British Columbia. For.Chron., 30:380-400.

59. Koot, P. 1994, Western hemlock looper. Natural Resources Can., Can. For. Serv., Pacific and Yukon Region, For. Pest Leaflet 21, 4pp. (A revision of Jardine 1969, and Erickson, 1984)

60. Krajina, V.J. 1965, Biogeoclimatic zones and classification of British Columbia. University of British Columbia Department of Botany, Vancouver, BC., pp. 1-17.

61. Krannitz, S.H. 1992, A synthetic sex pheromone for the western hemlock looper, *Lambdina fiscellaria lugubrosa* (Hulst) (Lepidoptera: Geometridae). M.Sc. Thesis, Simon Fraser Univ., Dept. Biol. Sci., Burnaby BC, 48pp.

62. Krieg, A. and Langenbruch, G.A. 1981, Susceptibility of arthropod species to *Bacillus thuringiensis*. In: Microbial Control of Pests and Plant Diseases 1970-1980. H.D. Burges, ed. Academic press., New York, NY, (pp. 837-896.)

63. Laflamme-Levesque, M., Perron, J.M., and Jobin, L. 1979, Study of the lichens belonging to the species *Bryoria Alectoria* and *Ramalina* on Anticosti Island related to *Lambdina fiscellaria fiscellaria* infestations. Can.Naturalist, 105:505-510.

64. Lertzman, K., Spies, T., and Swanson, F. 1997, From ecosytem dynamics to ecosystem management. In: The Rainforests of Home: Profile of a North American Bioregion. P.K. Schoonmaker, B. von Hagen, and S. Harrison, eds. Island Press, Covelo, California, (pp. 361-382.)

65. Li, W., Wang, Z., Ma, Z., and Tang, H. 1997, A regression model for the spatial distribution of red-crown crane in Yancheng Bioshpere Reserve, China. Ecological Modelling, 103:115-121.

66. Liang, Q., Otvos, I.S., and Bradfield, G.E. 1996, Distribution and sampling of eggs of the western hemlock looper (Lepidoptera: Geometridae) in mature western hemlock stands. J.Economic Entomology, 89:1531-1536.

67. Liang, Q., Otvos, I.S., and Bradfield, G.E. 1997, Forest roadside sampling of larvae and adults of the western hemlock looper, *Lambdina fiscellaria lugubrosa*. For.Ecol.Mgmt., 93:45-53.

68. Liang, Q., Otvos, I.S., and Bradfield, G.E. 1998, Pupal sampling of the western hemlock looper, *Lambdina fiscellaria lugubrosa* (Hulst.) (Lepidoptera Geometridae). J.Applied Entomology, 122:85-88.

69. Liebhold, A.M. and McManus, M.J. 1991, Does larval dispersal cause the spread of gypsy moth outbreaks? Northern Jounal of Applied Forestry, 8:95-98.

70. Liebhold, A.M., Zhang, X., Hohn, M.E., Elkington, J.S., Ticehurst, M., Benzon, G.L., and Campbell, R.W. 1991, Geostatistical analysis of gypsy moth (Lepidoptera: Lymantriidae) egg mass populations. Environ.Entomol., 20:1407-1417.

71. Liebhold, A.M., Elmes, G.A., Halverson, J.A., and Raske, A.G. 1994, Landscape characterization of forest susceptibility to gypsy moth defoliation. For.Sci., 40:18-29.

72. Liebhold, A.M., Gottschalk, K.W., Mason, D.A., and Bush, R.R. 1997, Evaluation of forest susceptibility to the gypsy moth across the conterminous United States. J.For., 95:20-24.

73. Liebhold, A.M., Gottschalk, K.W., Douglas, A.M., and Bush, R.R. 1997, Forest susceptibility to the gypsy moth. J.For., 43:483-490.

74. Lucarotti, C.J., LeClerc B., Morin B., and Ketella E.G. 1994, Development of an intergrated pest management program for the hemlock looper, *Lambdina fiscellaria fiscellaria*. Natural Resources Can., Can. For. Serv., Sci. Sustainable Development Dir., Ottawa, ON, (732pp.) 15-16

75. Mackey, B.G., McKenney, D.W., Yang, Y.Q., McMahon, J.P., and Hutchinson, M.F. 1996, Site regions revisited: a climatic analysis of Hill's site regions for the province of Ontario using a parametrc method. Can.J.For.Res., 26:333-354.

76. MacLean, D.A., Porter, K.B., Mackinnon, W.E., and Beaton, K.P. 2000, Spruce budworm decision support system: lessons learned in development and implementation. Computer and Electronics in Agriculture, 27:293-314.

77. Martinet, P.J. 1997, The role of climatic variation and weather in forest insect outbreaks. In: Insect Outbreaks. Academic Press Inc, (pp. 241-268.)

78. McDonald, D.M. and Nolan, R.A. 1995, Effect of relative humidity and temperature on *Entomophaga aulicae* conidium discharge from infected eastern hemlock looper larvae and subsequent conidium development. J.Invert.Path., 65:83-90.

79. McGuffin, W.C. 1987, Guide to the Geometridae of Canada (Lepidoptera) II Subfamily Ennominae. 4. Mem.Ent.Soc.Can., 138:1-182.

80. Medley, R.D. and Carolin, V.M.Jr. 1977, *Chaetophlepsis masellensis,* a dipterous parasitoid of the western hemlock looper (Lepidoptera: Geometridae). Can.Entomol., 109:1121-1128.

81. Montgomery, M.E. 1991, Variation in the suitability of tree species for gypsy moth. USDA Forestry General Technical Report NE-146 1-13

82. Morris, O.N. 1971, The effect of sunlight, ultraviolet and gamma radiations and temperature on the infectivity of a nuclear polyhedrosis virus. J.Invert.Path., 17:292-294.

83. Morris, O.N. 1983, Microorganisms isolated from forest insects of British Columbia. J.Entomol.Soc.Brit.Columbia, 80:29-36.

84. Norusis, M.J. 1994, SPSS Advanced Statistics 6.1. SPSS Inc., Chicago, Illinois, pp. 1-606.

85. Odland, J. 1988, Spatial Autocorrelation. Sage Publications, Newbury Park Beverly Hills London Delhi, pp. 1-85.

86. Ostaff, D.P., Borden, J.H., and Shepherd, R.H. 1974 Reproductive Biology of *Lambdina fiscellaria lugubrosa* (Lepidoptera: Geometridae). Can.Entomol., 106:659-665.

87. Otvos, I.S. 1967, The weather and hemlock looper numbers. Dept. Environ. Can. For. Serv., Newfoundland For. Res. Centre, Woody Points 7 (2) 8

88. Otvos, I.S., MacLeod, D.M., and Tyrell, D. 1973, Two species of *Entomophthora* pathogenic to the eastern hemlock looper (Lepidoptera: Geometridae) in Newfoundland. Can.Entomol., 105:1435-1441.

89. Otvos, I.S. 1977, Weather and outbreaks of the eastern hemlock looper in Newfoundland. Fish. Environ. Can., Can. For. Serv. Bi-mon. Res. Notes 33:3

90. Ouellet, M.J., Laflamme, J., and Perron, J.M. 1983, Effect of sugars on feeding behaviour of *Lambdina fiscellaria*. Ent.Exp.Appl., 34:139-142.

91. Packee, E. C. 1976, **The ecology of western hemlock**. Seattle University of Washington, College of Forest Resources,10-25

92. Parfett, N., Otvos I.S., and Van Sickle A. 1995, Historical western hemlock looper outbreaks in British Columbia: input and analysis using a geographical information system. FRDA II, FRDA report No. 235 1-36

93. Pereira, J.M.C. and Itami, R.M. 1991, GIS-based habitat modeling using logistic multiple regression: a case study of the Mt. Graham red squirrel. Photogram.Eng.Remote Sensing, 57:1475-1486.

94. Pojar, J., Klinka, K, and Meidinger, D. V. 1987, Biogeoclimatic ecosystem classification in British Columbia. For.Ecol.Mange 22, 119-154.

95. Raske, A.G. and Farrell J.E. 1986, Aspects of hemlock looper moth behaviour. Can. For. Serv. Newfoundland For. Centre., File Rpt. Study 2717 1-30

96. Roland, J., Mackey, B.G., and Cooke, B. 1998, Effects of climate and forest structure on duration of forest tent caterpillar outbreaks across central Ontario, Canada. Can.Entomol., 130:703-714.

97. Roland, J. 2001, Large-scale forest fragmentation increases the duration of tent caterpillar outbreak. Oecologia, 93:25-30.

98. Rothman, L.D. and Roland, J. 1998, Forest fragmentation and colony performance of forest tent caterpillar. Ecography, 21:383-391.

99. Ryan, R.B. 1974,Reduced oviposition by *Ephialtes ontario* and *Itoplectis quadricingulatus* in a humid environment. Annals of Ent.Soc.America, 67:928-930.

100. Safranyik, L., Shrimpton D.M., and Whitney H.S. 1980, Management of lodgepole pine to reduce losses from the mountain pine beetle. 1 Canadian Forestry Service, Victoria, B.C. 1-24

101. Sharov, A.A., Liebhold, A.M., and Roberts, E.A. 1997, Correlation of Counts of Gypsy Moths (Lepidoptera: Lymantriidae) in Pheromone Traps with Landscape Characteristics. For.Sci., 43:483-490.

102. Sharov, A.A., Pijanowski, B.C., Liebhold, A.M., and Gage, S.H. 1999, What effects the rate of gypsy moth (Lepidoptera: Lymantirriae) spread: winter temperature or forest susceptibility? Agric.and For.Entomology, 1:37-45.

103. Shepherd, R.H. 1994, Management strategies for forest insect defoliators in British Columbia. For.Ecol.Mgmt., 68:303-324.

104. Shore, T.L. and Safranyik L. 1990, A system for rating the susceptibililty and risk of lodgepole pine stands to mountain pine beetle. For. Can. PFC-File Rep. 5203 Forestry Canada, Victoria, BC 1-17

105. Shore, T.L. and Safranyik L. 1992, Susceptibility and risk rating system for the mountain pine beetle in lodgepole pine stands. BC-X-336. Forestry Canada, Victoria, B.C.

106. Silver, G.T. 1962, A review of some forest insect survey records associated with defoliator infestations in coastal British Columbia. Can. Dept. For., For. Ent. Path. Branch, For. Ent. Path. Laboratory, Victoria, BC, Interim Rpt., 23 pp. (unpublished)

107. Smirnoff, W.A. and Jobin, L. 1973, Study of certain factors affecting populations of *Lambdina fiscellaria fiscellaria* in the watershed of the Vaureal River of Anticosti Island. Can.Entomol., 105:1039-1040.

108. Taylor, R.J. 1993, Flora of North America North of Mexico. Oxford University Press,

109. Thompson, M.G. 1956, Studies of factors affection reproduction in the western hemlock looper, *Lambdina fiscellaria lugubrosa* (Hulst) (Lepidoptera: Geometridae). Can. Dept. Agric., Scie. Serv., For. Biol. Laboratory, Victoria BC, Interim Rpt. 1955-4, 41 pp.

110. Thompson, M.G. 1958, The estimation of larval forest defoliator populations by frass drop and body weight. Can. Dept. Agric, Sci. Serv., For. Biol. Laboratory, Victoria, BC. 1-14

111. Thompson, M.G. 1958, Egg sampling for the western hemlock looper. For.Chron., 34:248-256.

112. Torgersen, T.R. 1971, Parasites of the western hemlock looper, *Lambdina fiscellaria lugubrosa*(Hulst) in southeast Alaska. The Pan-Pacific Entomologist, 47:215-219.

113. Trani, M.K. and Giles, R.H.Jr. 1999, An analysis of deforestation: Metrics used to describe pattern change. For.Ecol.Mgmt., 114:459-470.

114. Turnquist, R. 1991, Western hemlock looper in British Columbia. FIDS Rpt. 91-8 For. Can., Pacific For. Centre, Victoria 1-42

115. Van Sickle, G.A. 1994, Status of important forest pests and experimental control projects in the Pacific and Yukon Region in 1994. Rpt. of the Twenty-second Ann. For. Pest Control Forum, Nov. 15-17. 1994, 313 pp. Natural Resources Can., Can. For. Serv., Ottawa, ON 27-43

116. Van Sickle, G.A. 1995, Forest insect Pests in the Pacific and Yukon Region. In: Forest Insect Pests in Canada (388pp.). J.A. Armstrong and W.G.H. Ives, eds. Natural Resources Can., Can. For. Serv., Sci. Sustainable Development Dir., Ottawa, ON, (pp. 73-89.)

117. Van Sickle, G.A., Fiddick, R.L., and Wood, C.S. 2001, The Forest Insect and Disease Survey in the Pacific Region. J.Entomol.Soc.Brit.Columbia, 98:169-176.

118. Virtanen, T., Neuvonon, S., and Nikula, A. 1988, Modeling topoclimatic patterns of egg mortality of *Errita autumnata* (Lepidoptera: Geometridae) with a geographical information system: predictions for current climate and warmer climate scenarios. Journal of Applied Ecology, 35:311-322.

119. Wang, W.L.Z. and Tang, Z.M.H. 1997, A regression model for the spatial distribution of red-crown crane in Yancheng Biosphere Reserve, China. Ecological Modelling, 103:115-121.

120. Weseloh, R.M. 1996, Developing and validating a model for predicting gypsy moth (Lepidoptera: Lymantriidae) defoliation in Connecticut. Forest Entomology, 22:1546-1555.

121. Wyatt, G.R. 1946, Hemlock looper-Lake Cowichan field studies 1944-45. Can. Dept. Agric., Scie. Serv.. For. Biol. Laboratory, Interim Tech. Rpt. Victoria BC 1-14

122. Yang, D., Pijanowski, B.C., and Gage, S.H. 1998, Analysis of gypsy moth (Lepidoptera: Lymantriidae) population dynamics in Michigan using geographic information systems. Environ.Entomol., 27:842-852.

123. Stark, R.W. 1982, Bark beetles in North American conifers: a system for the study of evolutionary biology. University of Texas Press, Austin, Texas, pp. 1-557.

Appendix 1. Percentage of total western hemlock looper defoliation (1911-1995) by Biogeoclimatic sub-zones

Biogeoclimatic Zones		
Zone	**Subzone**	**% of defoliated Area**
AT (Alpine Tundra)	AT	0.6%
CWH (Coastal Western Hemlock)	CWH dm	0.6%
	CWH mm	0.1%
	CWH vh	0.0%
	CWH vm	3.1%
	CWH xm	0.1%
ESSF (Englemann Spruce-Sub Alpine Fir)	ESSFmm	0.0%
	ESSFvc	3.3%
	ESSFwc	6.0%
	ESSFwcp	0.0%
	ESSFwk	1.0%
	ESSFwm	0.1%
ICH (Interior Cedar Hemlock)	ICH dw	0.0%
	ICH mk	0.0%
	ICH mw	21.0%
	ICH mm	0.4%
	ICH vk	20.4%
	ICH wk	38.6%
IDF (Interior Douglas-fir)	IDF dk	0.0%
	IDF mw	0.0%
	IDF xh	0.5%
MH	MH mm	0.0%
SBS (Sub-Boreal Spruce)	SBS dh	0.1%
	SBS vk	3.6%
	SBS wk	0.3%
		100.0%

Appendix 2. Projection parameters

Lambert conformal conic projections maintain accurate local shape and direction. They are best suited for mid latitude applications and east-west oriented maps. Geographers recommend using them for continental and regional applications where total range in latitude does not exceed 35 degrees. This projection is a secant projection whereby the two standard parallels pass through the sphere and areal distortion is minimized at the contact points. Scale is correct at the parallels as are local angles. Projection parameters for the maps and analysis in this study are as follows:

Projection: Lambert
Units: Metres
1st Standard Parallel: 51° 00’ 00”
2nd Standard Parallel: 58° 00’ 00”
Central Meridian: -126° 00’ 00”
Latitude of projection origin: 48° 00’ 00”
False easting (metres): 1 000 000
False northing (metres): 0

Appendix 3. Measurement accuracy and precision standards for FIP/FC-1 Data 1996

Accuracy and Measurement Standards

Attribute	Standard
Species Composition	Correct for first two species ≥ 6 % by composition
Age	± 20 % of the most representative age for leading species
Height	>20m: ± 20% of the most representative height for leading tree species >20m: ± 10% of the most representative height for the leading species
Crown closure	± 10% of the most representative corwn closure for trees in the main canopy
Site Index	For stands > 20 years, the class should represent the growth potential of the site based on the leading species.

(modified from BCMOF, 1992)

Precision of Estimates:

Attribute	When Practical to the Nearest	Otherwise to the Nearest
Species Composition	1%	10%
Age	1 year	10 years
Height	0.1 m	1 m
Crown closure	1%	10%
Density	1 tree	10 or 100 trees
Reference Year*	1 year	-

(modified from BCMOF, 1992)

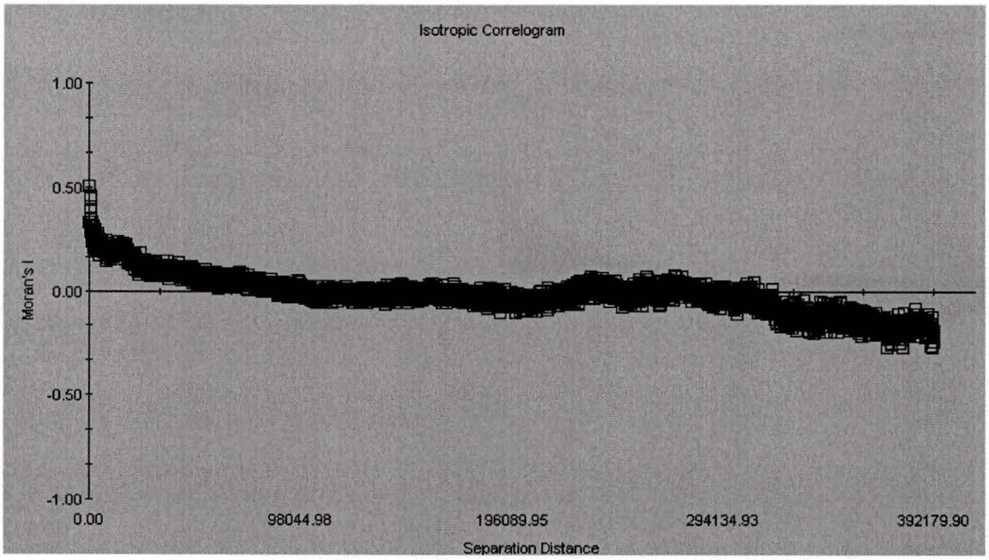
*Reference year, although not included in the actual statistical analysis, is a constraint in the selection of forest polygons, and all polygons are current as of 1992. This refers to the calendar year for which the forest type attributes are recorded and most accurate. Age and height current references are the basis for this attribute.

It is apparent from a frequency examination of the variables in the dataset that the majority of the forest type attributes fall into the larger bin sizes, such as 10%/years/trees etc. Smaller units are recorded with less frequency and, therefore, the resolution of the dataset ought to be considered to be analyzed at the coarser frequency. As a rule of thumb, the younger the stand's age, the finer the resolution of the data, presumably due to the presence of silviculture records (versus tree-coring and air-photo interpretation for older stands).

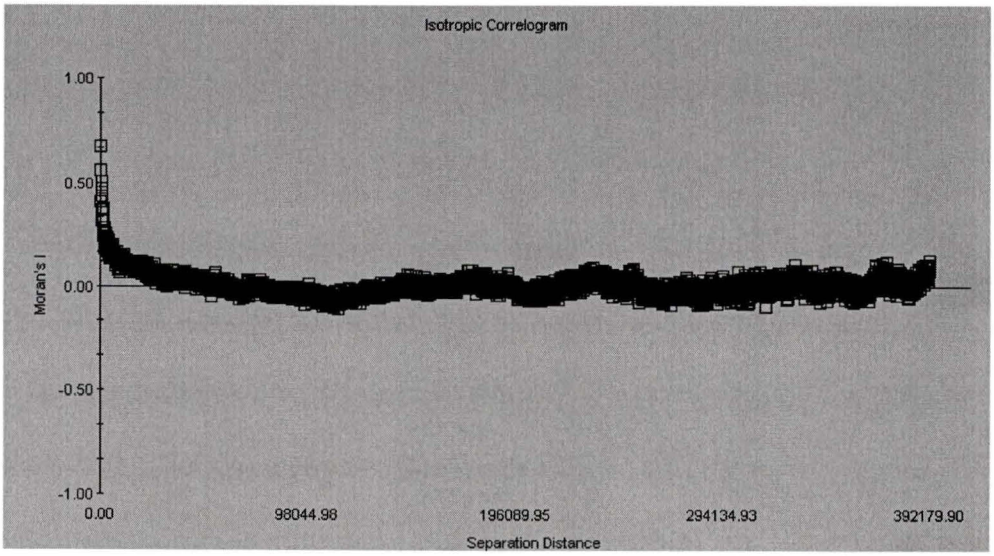
Appendix 4. Autocorrelation and isotropic correlograms

The following appendix presents the isotropic correlograms for the variables exhibiting the most spatial autocorrelation. These correlograms are based upon Moran's I-statistic (analagous to Pearson's r-statistic), but interpreted for the values of a forest stand variable at distances (on the x-axis) between the centroids of forest stands.

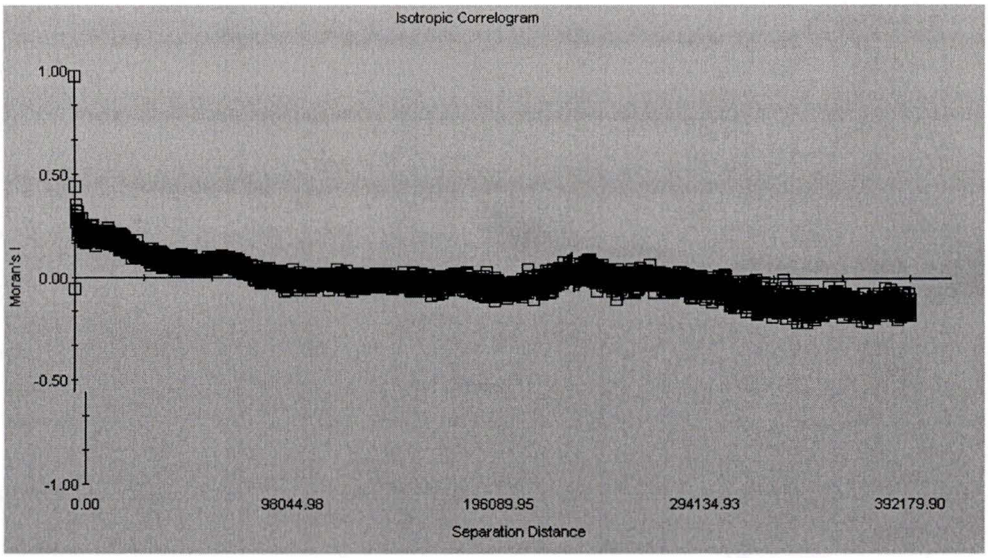
Site Index



Hemlock and Cedar Content

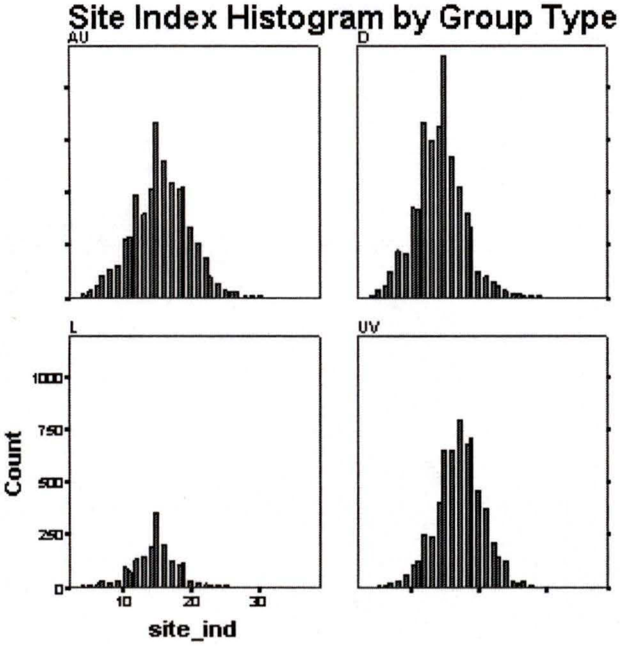


Stand Age

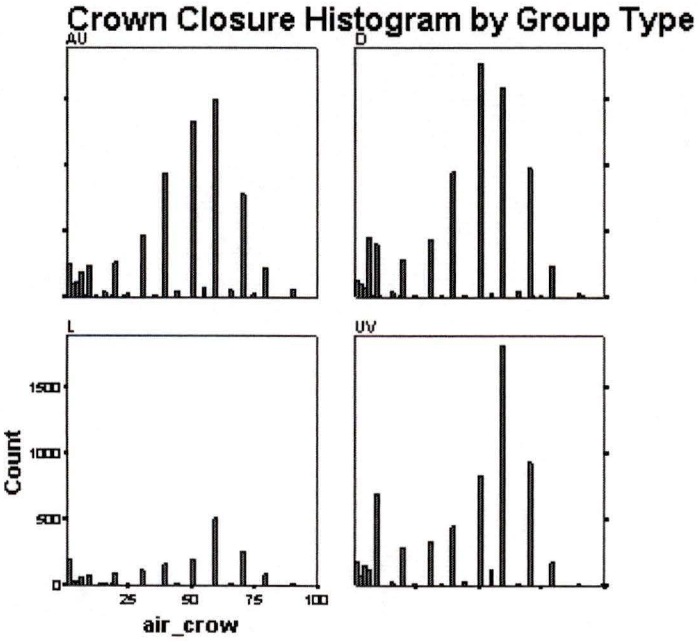

Isotropic Correlogram
1.00
0.50
0.00
-0.50
-1.00
Moran's I
0.00
98044.98
196089.95
294134.93
392179.90
Separation Distance

Appendix 5. Stand variable histograms

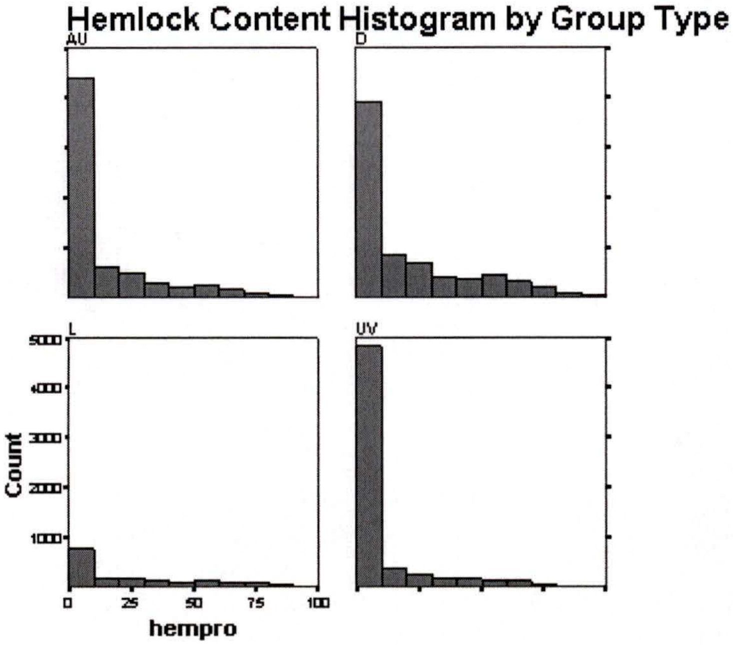
Site Index



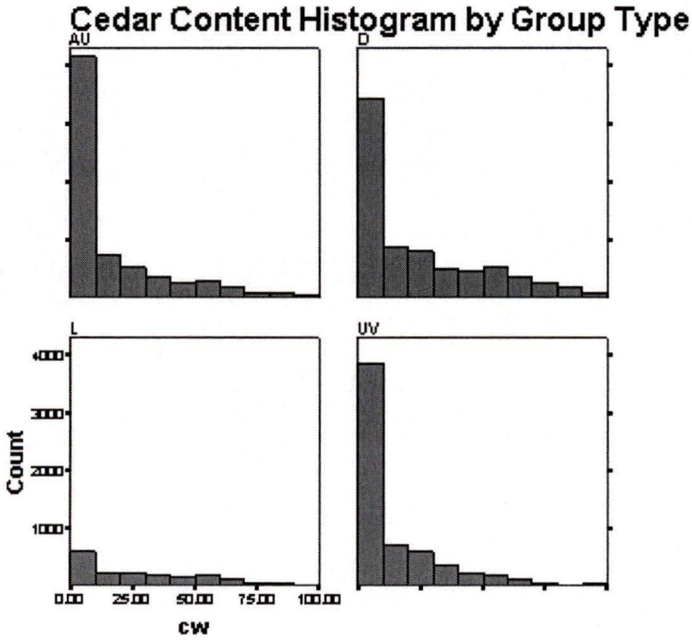
Crown Closure



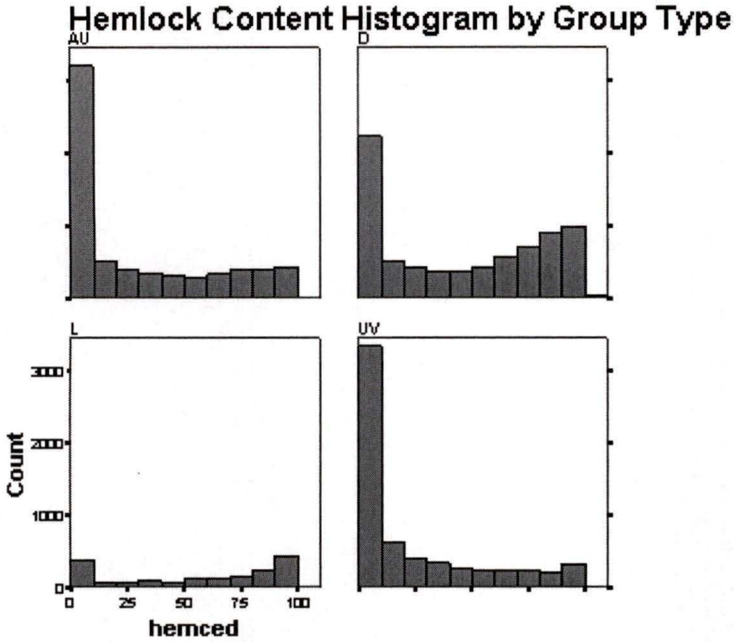
Hemlock Content



Cedar Content



Hemlock and Cedar Content


Hemlock Content Histogram by Group Type

Vita

Surname: Borecky Given Names: Neil

Place of Birth: Richmond Hill, Ontario, Canada

Educational Institutions Attended:

McMaster University	1992 to 1996

Degrees Awarded:

B.A. (Honours)	McMaster University	1996

Honours and Awards:

Science Council of British Columbia GREAT Scholarship	1999 to 2000

Publications:

Borecky, Neil and Otvos, Imre S. 2001, Coarse scale hazard rating of western hemlock looper (*Lambdina fiscellaria lugubrosa*) in British Columbia. *In:* Proceedings integrated management and dynamics of forest defoliatin insects, 1999, United States Department of Agriculture, Penn. pp.6-15

UNIVERSITY OF VICTORIA PARTIAL COPYRIGHT LICENSE

Title of Thesis:

Development of a western hemlock looper *Lambdina fiscellaria lugubrosa* hazard rating system for Interior British Columbia using discriminant function analysis and logistic regression.

Author [redacted]

[redacted]

Neil Borecky

August 28, 2003