

**An Assessment of the Value of Landsat Imagery
for the Most Similar Neighbour Procedure:
West Arm Demonstration Forest Case Study**

**An Assessment of the Value of Landsat Imagery
for the Most Similar Neighbour Procedure:
West Arm Demonstration Forest Case Study**

Prepared for

BC Ministry of Forests
Forest Practices Branch
P.O. Box 9513, Stn. Prov. Gov't
Victoria, BC V8W 9C2

Prepared by

Donald C.E. Robinson
ESSA Technologies Ltd.
1765 West 8th Avenue, Suite 300
Vancouver, BC V6J 5C6

September 30, 1999

Citation: **Robinson, D.C.E.** 1999. An Assessment of the Value of Landsat Imagery for the Most Similar Neighbour Procedure: West Arm Demonstration Forest Case Study. Prepared by ESSA Technologies Ltd., Vancouver, BC for BC Ministry of Forests, Victoria, BC. 14 pp.

© 1999 ESSA Technologies Ltd. and BC Ministry of Forests

No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without prior written permission.

Acknowledgements

We express our appreciation to Graham Smith of Geosense Consulting, Nelson BC, for quickly providing us with Landsat imagery and for arranging for its use with the MELP office in Nelson. We also extend our thanks to Melinda Moeur of the USDA Forest Service Rocky Mountain Research Station, Moscow ID, for helpful discussions relating to the use of matches and mismatches, and for discussion of the role of tie-breaking in the assignment of nearest neighbours.

This report was formatted and produced by Kelly Robson.

Funding for this research was provided by Forest Renewal BC — a partnership of forest companies, workers, environmental groups, First Nations, communities and government. Forest Renewal BC funding — from stumpage fees and royalties that forest companies pay for the right to harvest timber on Crown lands — is reinvested in the forests, forest workers, and forest communities. Funding assistance by Forest Renewal BC does not imply any endorsement of any statements or information contained herein.

Table of Contents

List of Tables	iv
List of Figures.....	iv
1.0 Introduction.....	1
1.1 <i>Background</i>	1
1.2 <i>Objectives</i>	2
2.0 Methods.....	3
2.1 <i>Study Location</i>	3
2.2 <i>Acquisition of Low Resolution Indicator Data</i>	3
2.3 <i>Acquisition of High Resolution Design Data</i>	4
2.4 <i>Experimental Framework</i>	6
3.0 Results & Discussion.....	7
3.1 <i>All-Data Analysis</i>	7
3.2 <i>Effect of Ties</i>	7
3.3 <i>Effect of Outliers</i>	8
3.4 <i>Classified Versus Raw Landsat Data</i>	8
4.0 Discussion & Conclusions.....	11
4.1 <i>Required Sampling Effort</i>	11
4.2 <i>Value of Landsat Data</i>	11
5.0 Literature.....	13

List of Tables

Table 1.1:	The goal of the MSN procedure	1
Table 2.1:	A summary of MSN analyses, leaving out all 6 combinations of the 3 data classes	7
Table 2.2:	A summary of MSN analyses using raw or clustered Landsat data, leaving out all 6 combinations of 3 data classes.....	9

List of Figures

Figure 2.1:	The Redfish Creek Study Area, located within the West Arm Demonstration Forest near Nelson, BC	5
--------------------	--	---

1.0 Introduction

1.1 Background

Forest resource managers and planners are regularly called upon to design and evaluate land use plans for landscapes of thousands of acres, comprising stands with diverse composition and structure and possibly spanning multiple ecosystems. To measure the impacts of different land use scenarios, these plans must forecast future conditions. They therefore require knowledge of the current state of each unique parcel of land, in order to drive the simulation models that project the future state of the stand.

Some planning models may have substantial data requirements. For example, the Prognosis stand simulation model (Wykoff et al. 1982) requires a statistical sample of the trees in each stand in order to project future development and structure. The Prognosis Environmental Indicators model (Greenough et al. 1999a), which incorporates the BC variant of Prognosis and three other extensions, requires additional estimates of coarse woody debris for each stand, in addition to the sample of living and recently killed trees.

Collecting the detailed field information needed for the projection of each parcel is not feasible for such complex models. The Most Similar Neighbour (MSN; Moeur and Stage 1995) procedure provides a statistically based method for inferring the detailed field information for unsampled stands based on the (lower-resolution) data that are available for both sampled and unsampled stands. Such broad coverage data are commonly available through digital elevation models and from the analysis of aerial photography. MSN uses the relationship between the known characteristics of sampled and unsampled stands – as quantified by canonical correlation – to calculate a unit-independent multivariate measure of distance between all pairs of stands. Those stands that have the smallest distance are the most similar. Matching each stand to the most similar sampled stand then imputes detailed data for each unsampled stand. Table 1.1 illustrates the MSN procedure conceptually. With reference to this table, the goal of the procedure is to find the most appropriate data to fill in the shaded lower-left portion of the table.

Table 1.1: The goal of the MSN procedure. High-resolution data are available only for sampled stands, not the unsampled stands shown in the shaded box of the figure. The goal of the MSN procedure is to provide the best high-resolution data for the low-resolution stands by matching each low-resolution stand with the most suitable stand for which high-resolution data are available. This is done by using the low-resolution data available for both sampled and unsampled stands (the right-hand boxes of the figure) to identify the most similar sampled stand in the upper left box. See text for further explanation.

High resolution Design data for sampled stands: 1. 2. 3. ...	Low resolution Indicator data for sampled stands: 1. 2. 3. ...
<i>Inferred</i> high resolution Design data for unsampled stands: 1. 2. 3. 4. ...	Low resolution Indicator data for unsampled stands: 1. 2. 3. 4. ...

Compared to alternative procedures such as stratified sampling, the chief advantages of the MSN procedure are that it both

- faithfully preserves the range of variation found in the original data; and
- retains the correlation structure found in the original data.

1.2 Objectives

Three classes of lower resolution Indicator data are usually available in British Columbia:

- topographic data in the form of a 1:20,000 digital elevation model;
- current FIP projections (photographically based with ground validation and updated by VDYP); and
- remotely-sensed Landsat spectral data.

A recent application of the MSN (Greenough et al. 1999b; details in Section 2) was able to make use of the first two kinds of data: topographic and FIP. Landsat imagery for the same study area has now been obtained, and it is now possible to compare the relative contributions of all three classes of data to MSN analyses. In MSN applications that may have limited financial resources for inventory, this comparison may lend some insight into the most effective use of those resources.

The objective of this study is therefore to quantify the value of satellite imagery to the MSN procedure, comparing its contribution to the use of topographic and FIP data alone.

2.0 Methods

2.1 Study Location

This comparison of the contributions to MSN analyses of Landsat imagery, topographic information and the provincial inventory database is based on data from the Redfish Creek watershed; a 3,300 ha landscape within the West Arm Demonstration Forest near Nelson, BC. As described in detail below, the Redfish Creek watershed has recently been intensively sampled as part of the development of the Prognosis Environmental Indicators extension (Greenough et al. 1999b). The intensive effort already made to gather and compile new and existing information in this watershed greatly simplifies the addition of recently acquired spectral data.

2.2 Acquisition of Low Resolution Indicator Data

The BC Ministry of Forests 1:20,000 digital Forest Cover map and its associated FIP data are the basis for most of the Indicator data. In the Redfish Creek study area there are 152 polygons and FIP records. These forest cover polygons have been further subdivided, based on soils mapping (Utzig 1997), to define 496 ecologically-based polygons. From the FIP data, all hardwoods (birch and aspen) have been combined and one instance of whitebark pine classified with lodgepole pine. Slope, aspect and elevation are based on area-weighted mean values resulting from the overlay of the digital elevation model themes from the Forest Cover maps onto the 496 polygons defined for the study area (see Figure 2.1). Timber volume estimates of each species in each polygon are based on VDYP projections to 1997, including age, height and volume (m³/ha, 12.5 cm lower dbh limit).

The MOF data for the polygons include the 21 variables (and transformations) shown below. Of these variables, 16 are for forest cover-oriented FIP data and 5 are for topographically oriented site data:

inventory crown cover	Se (Engelmann spruce) volume
age	Bl (subalpine fir) volume
height (m)	Pp (ponderosa pine) volume
height ²	hardwood volume
Pw (western white pine) volume (m ³ /ha)	site index
Lw (western larch) volume	slope (%)
Fd (Douglas-fir) volume	cosine aspect (radians)
Bg (grand fir) volume	sine aspect
Hw (western hemlock) volume	cosine aspect \times slope
Cw (western redcedar) volume	elevation (m)
Pl (lodgepole pine) volume	

The third class of Indicator data are based on Landsat imagery made August 10, 1997.¹ The Landsat Thematic Mapper data are utilised in two ways. The first method uses the raw values of spectral bands 3, 4 and 5, expressed as integer values from 0-255 on 25-m resolution Arcview grids. The grids were overlaid on the 496 polygon Study Area, and the mean and variance of each band computed for each polygon. The second method uses an unsupervised classification procedure² to process the same

¹ Landsat 5, track 44, frame 25.

² PCI Image Analysis Software Suite (version 6.3), ImageWorks Module; Procedure ISOCLUS.

information in the 3 spectral bands. This classification produced 11 cluster-types within the Study Area. The map of cluster-types was overlaid on the 496 polygons and the proportion of each polygon's area occupied by each cluster type was reported and coded as the proportion of the cluster type in the polygon.

The Landsat data for the polygons include the following 7 variables and 2 transformations. The transformations follow those suggested by Moeur and Riemann Hershey (1999), where Normalised Difference Vegetation Index (NDVI) is based on Reed et al. (1994).

Raw Data

mean band 3
 mean band 4
 mean band 5
 variance band 3
 variance band 4
 variance band 5
 NDVI: $(\text{band 4} - \text{band 3}) / (\text{band 4} + \text{band 3})$
 band 3 / band 4

Clustered data

proportion area in clusters 1-11

2.3 Acquisition of High Resolution Design Data

As part of a study to develop and test the Prognosis Environmental Indicators extension (Greenough et al. 1999b), 69 plots were established in the Redfish Creek Study Area by Katherine Wolters (Surewood Forest Consultants, Nelson BC) and Barry Snowdon (BC Ministry of Forests). The final number of "Surewood plots" was constrained by the available budget and by the length of the sampling season. Plot locations were determined by laying out a randomly aligned 500-m square grid on a map of the watershed, insuring that all major ecological zones were sampled. Each sample point was measured by fixed area plot. Samples included increment cores, heights and live crown length for the major species and a statistical sample of all live trees.

A further 15 plots ("SRC plots") were available from the Prognosis EI case study. These plots were inventoried during the preparation of an earlier logging plan (J. Hammons, Simons Reid Collins, pers. comm.), and were based on prism plots rather than fixed area plots. An additional group of SRC plots were not applicable in the present analysis, because harvests in those locations made the current Landsat data incomparable to the pre-harvest state of the SRC inventories.

The Surewood and SRC plots were converted to Prognosis treelists and summarised for the year 1997 using the SEI variant of Prognosis (ver. SI 1.0i; Robinson 1998). An estimate of each species' volume was made by post-processing the output treelist for 1997. A merchantable volume (m^3/ha) diameter limit of 12.5 cm dbh was chosen to make the volumes comparable with volume estimates from the FIP data. Slope, aspect and elevation data from the sampled stands were also incorporated. The data in the upper left corner of Table 1.1 therefore included 12 forest cover (FIP) variables and 5 Topographic variables:

Pw (western white pine) volume (m^3/ha ; 12.5 cm dbh)	Pp (ponderosa pine) volume
Lw (western larch) volume	hardwood volume
Fd (Douglas-fir) volume	CCF
Bg (grand fir) volume	slope (%)
Hw (western hemlock) volume	cosine aspect (radians)
Cw (western redcedar) volume	sine aspect
Pl (lodgepole pine) volume	cosine aspect \times slope
Se (Engelmann spruce) volume	elevation (m)
Bl (subalpine fir) volume	

To find the spatial correspondence between the Design data locations and the Indicator data locations (polygons), the Surewood plot locations were digitised over the FIP theme, using paper maps of the sampling grid (Figure 2.1). Locations for the SRC plots were derived from annotations on the cruise cards. These annotations provided provincial Forest Cover polygon numbers and road and stream drawings made in the field. We believe that the digitised location of the plots is within 50 m of the true plot location. The high resolution inventories included estimates of slope, aspect, elevation and BEC/Site-Series (BEC/ss) taken at the time of sampling. BEC/ss designations made during the inventory collection were compared with those developed by Ketcheson (1992, 1993) and found to be in agreement in all cases. The corresponding FIP information at each sample point was used to fill in the data for the lower resolution portion of the data (*i.e.*, completing the upper right corner of Table 1.1).

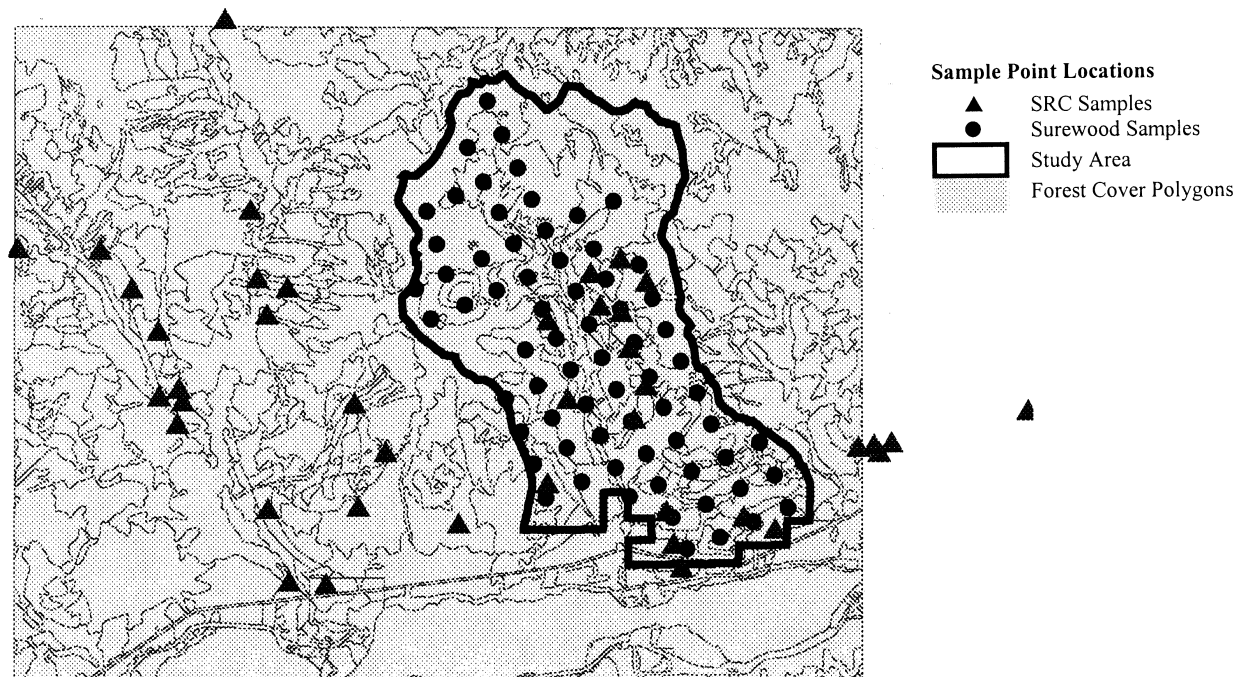


Figure 2.1: The Redfish Creek Study Area, located within the West Arm Demonstration Forest near Nelson, BC. Within the study area, the map shows 152 Forest Cover polygons. As described in the text, these polygons were further subdivided into 496 polygons based on soils mapping. Only those points located within the Study Area were used in the current MSN analysis.

2.4 Experimental Framework

The high- and low-resolution data were prepared and formatted for the MSN programme³ using a combination of custom Arcview Avenue-scripts for the data preparation steps, and custom Korn Shell scripts for the analysis.

Three experiments were conducted. The first selectively removed each of the three classes of low-resolution Indicator data (Topographic, FIP and Landsat) one-at-a-time, so that the effect of each data class on the most similar neighbour assignment could be compared to the assignment made using the all three data classes. Interactions were then examined by removing the Indicator classes two-at-a-time.

As explained already, the MSN procedure computes a multivariate “distance” between each high-resolution and low-resolution stand. An initial examination of an MSN analysis using all observations and all classes of data found 15 low-resolution stands with very high distances. The second analysis therefore examined the effect of excluding these high-distance outliers and the sensitivity of the data comparison to the presence or absence of outliers.

Finally, most-similar-neighbour assignments were made using raw Landsat data (mean and variance of the spectral bands), and compared to the assignments made using spectral data processed by a clustering algorithm.

The MSN assignments made using all the data were used as benchmarks for comparison in the analyses. That is, the most-similar-neighbour for each of the 496 polygons as determined with complete data from all 3 information classes was taken to be the best classification, and changes in that assignment were taken as a measure of reduced fit. This approach is consistent with analyses made by Moeur (pers. comm.).

³ MSN version 3.00, obtained by anonymous FTP from forest.moscowfsl.wsu.edu/trees/newmsn.zip and modified slightly to compile with MS Fortran PowerStation 4.0.

3.0 Results & Discussion

3.1 All-Data Analysis

In the first analysis, each of the 496 polygons was assigned one of the 84 inventory plots using all three classes of Indicator data: Topographic, FIP and Landsat. The inventory plot assignments made this way were treated as the benchmark assignments. Leaving out one and then two classes of data resulted in the reclassification of each stand using less auxiliary information. The results of this analysis are summarized in Table 2.1.

Table 2.1: A summary of MSN analyses, leaving out all 6 combinations of the 3 data classes. When all 3 data classes are included (top row), each of the 496 polygons is matched to itself, by definition. As data classes are removed from the canonical correlation, the MSN procedure may find different most similar neighbours, resulting in a change in assignment between the benchmark analysis and the analysis leaving data out. The number of **identical** assignments is shown here.

Number of Classes	Data Class			Identical Matches
	Topographic	FIP	Landsat	
3				496
2				273
				218
				257
1				84
				223
				151

The leave-two-classes-out (bottom three rows of Table 2.1) analysis clearly shows that when only one class of data is available, the FIP database (volume estimates, age, height, Site Index and Crown Cover) produces neighbour assignments most like those made when all data are available (223 of 496). Landsat alone and Topographic alone produce identical assignments in 30% and 17% of cases, compared to 45% for the FIP data. Note though, that the Topographic-alone agreement (84/496) is tantamount to no matches, since the 84 Design observations will always match the polygon they are found in, by definition.

When two sources of data are available, the FIP-Topographic and FIP-Landsat combinations are the most informative, improving the FIP-only assignment to about 53% agreement with the all-data assignment. Topographic-Landsat in combination fare about as well as FIP alone.

3.2 Effect of Ties

In about 15% of cases, an unsampled low-resolution polygon has more than one most similar neighbour. The MSN programme resolves these ties by using a random number to select one of the Design polygons at random. The programme does not currently allow users to control the random number, nor does it report the identities of multiple matches. The differences reported in Table 2.1 may therefore overstate the magnitude of different matches, since some 'differences' may actually be ties that are broken in favour of

a different (with respect to the benchmark) Design polygon. The ambiguity introduced by tie-breaks accounts for part of the large jump that occurs between the benchmark matches and the leave-one-out analyses, since by definition there is no tie-breaking ambiguity to affect the benchmark. This will affect the absolute value of the differences, but because the leave-out analyses are all affected in the same way, the results are only shifted (made larger) and their ranking is unaffected. Characterizing the importance of tie-breaks in the assignment of the most similar neighbour would help to clarify the statistical significance of the differences reported in Table 2.1. However, this is beyond the current capabilities of the MSN programme.

3.3 Effect of Outliers

The majority of the polygons are well matched by the MSN All-Data analysis: 97% of the polygons fall within the first decile of computed MSN distances, leaving only 15 outliers in the remaining 9 deciles. By studying other information in the FIP database, these outlier polygons were identified with features such as swamps, lakes, krummholz, alpine areas and powerlines. Only one polygon had what appeared to be truly anomalous species composition. To explore the impact these outliers might have, identical 'leave-one-out' and 'leave-two-out' analyses was made using the Indicator data without these outliers. Outlier removal did not affect the relative ordering shown in Table 2.1. However, the absolute numbers were lower in all cases due to the lower number of polygons in the landscape, and some other small shifts in difference were noted, possibly caused by tie-breaking.

3.4 Classified Versus Raw Landsat Data

The final set of analyses split the Landsat data into its two component parts:

- 11 classes defined by an unsupervised clustering algorithm; and
- raw spectral data with some transformations.

In each polygon the class data were coded in proportion to the area of each class. The raw data were represented as continuous variables measured by the mean and variance and by two transformations.

The All-Data analysis (Section 3.1) was repeated, retaining the original benchmark assignments for the comparison of new assignments. These results are summarized in Table 2.2.

Table 2.2: A summary of MSN analyses using raw or clustered Landsat data, leaving out all 6 combinations of 3 data classes. The original MSN analysis of Table 2.1 (using both classified and raw Landsat data) continues to serve as the benchmark for comparisons. The final two columns show the **change** in the number of identical assignments relative to Table 2.1.

Number of Classes	Data Class			Change in Identical Matches	
	Topographic	FIP	Landsat	Raw	Clustered
3				-192	-217
2				-24	-32
				-31	-15
				0	0
1				0	0
				0	0
				-26	-59

As anticipated, those MSN analyses in Table 2.2 that make no use of satellite data are unchanged with respect to Table 2.1. The other analyses show fairly consistent results, with raw data performing better than clustered data in 3 of four cases. When the Landsat data are used in concert with other data sources, the differences are not as pronounced as when the Landsat data are used alone (bottom row).

4.0 Discussion & Conclusions

4.1 Required Sampling Effort

Relative to other methods of ‘filling in’ missing data, the chief advantages of the MSN procedure are that it preserves the range of variation and the correlation structure found in the input data. One further question that arises is the amount of intensive sampling effort required for the valid imputation of missing data. For example, the sampling effort recently applied to the Prognosis EI case study in the West Arm Demonstration Forest (Greenough et al. 1999b) was able to inventory 23% of the 496 low resolution polygons. This issue is an ongoing area of study, and the answer depends in part upon the variability present within the landscape it may be applied to. A more homogeneous landscape will require a less intensive sampling effort compared to a heterogeneous landscape, but this fact may not be known prior to the inventory.

A recent study in the Deschutes National Forest (Moeur 1999) has used a completely sampled landscape of 2,000 stands to study this question more fully. By leaving out sampled stands and then imputing their attributes with MSN, the predictive abilities (including error and bias) of the procedure can be tested with different levels of simulated sampling effort. Briefly, the four main findings are that:

- bias (in % cover, basal area and volume) becomes stable and zero when more than 20% of the stands are sampled;
- root mean square error for these three attributes drops linearly to zero as 0-100% of stands are sampled;
- the average and maximum MSN ‘distance’, and the maximum number of times a stand is selected as the best neighbour, all decrease as more high-resolution stands are added; and
- graphs of these relationships all flatten out at between 20% and 30% of total stands sampled.

Taken together, the results for this landscape indicate that the incremental benefit of additional high-resolution Design stands declines when more than 20-30% of the landscape is sampled. The general validity of a 20% rule-of-thumb to other landscapes is not yet known, but the level of effort that has been applied in the Prognosis EI case study seems to be appropriate.

4.2 Value of Landsat Data

A number of straightforward conclusions emerge from this study of the role of satellite and conventional data in the Most Similar Neighbour procedure.

Using all available data – the so-called ‘kitchen sink’ approach (Moeur 1999) – as a benchmark, the single best source of MSN classification data is the FIP database. The expertise embodied in the assignment of polygons and cover from aerial photography, coupled with on-site estimation of site index and species composition, is the best primary source of information for the MSN procedure. When two sources are available, FIP in concert with spectral data are the most useful. Topographic information (slope, aspect and elevation) alone is not a good basis for assigning most similar neighbours, but is a useful source of auxiliary information.

Perhaps surprisingly, raw spectral data (the mean and variance of bands 3-5 with some transformations) are more useful than clustered data for MSN purposes. The information embodied in the correlation between raw spectral bands and their transforms is easier to obtain than clustered information – requiring one less step in data preparation – and better captures the relationships between the spectral bands that embody the characteristics of the underlying vegetation.

The value of the 3 kinds of information can be summarized:

FIP > Landsat > Topographic

This ranking reflects the value of each data type alone and also the order in which new data sources should be added. The large jump in agreement that is found when all 3 data types are included in the MSN analysis, compared to only 2 types, shows the benefit of including all data when data acquisition costs are not an issue.

Outliers and tied-assignments are two possible confounding features of the analysis. In this case study, the presence of outliers does not affect our conclusion about the relative value of the three sources of low-resolution Indicator data. The effect of tied-assignments also should not bias the ranking of the three data classes, although it does change the absolute number of matches and mismatches in ways that are not yet known. This uncertainty could easily be removed in future development of the MSN software by allowing users to control the tie-breaking process through the control of random number seeds.

5.0 Literature

Greenough, J.A., D.C.E. Robinson, S.J. Beukema, and W.A. Kurz. 1999a. Prognosis Environmental Indicators Model Description. Prepared by ESSA Technologies Ltd., Vancouver, BC. 66 pp.

Greenough, J.A., D.C.E. Robinson, S.J. Beukema, and W.A. Kurz. 1999b. Prognosis Environmental Indicators Model: WADF Case Study Methodology. Prepared by ESSA Technologies Ltd., Vancouver, BC. 24 pp.

Ketcheson, M.V. 1992. Preliminary biophysical habitat map legend: West Arm Demonstration Forest, Kootenay Lake Forest District. Review draft prepared by JMJ Holdings Inc., Nelson, BC. 72 pp.

Ketcheson, M.V. 1993. WADF biophysical habitat mapping project report. Prepared by JMJ Holdings Inc., Nelson, BC.

Moeur, M. and A.R. Stage. 1995. Most Similar Neighbor: an improved sampling inference procedure for natural resource planning. *For. Sci.* 41:337-359.

Moeur, M. 1999. Extending stand exam data with Most Similar Neighbor inference to update a current vegetation inventory for the Deschutes National Forest, Oregon. Unpublished mss on file at Rocky Mountain Research Station, Moscow, ID.

Moeur, M. and R. Riemann Hershey. 1999. Preserving spatial and attribute correlation in the interpolation of forest inventory data. In: Proceedings, 3rd international symposium on spatial accuracy assessment in natural resources and environmental sciences, May 20-22, 1998, Quebec City, Canada. Ann Arbor Press.

Reed, B., J.F. Brown, D. VanderZee, T.R. Loveland and J.W. Merchant. 1994. Measuring phenological variability from satellite. *Journal of Vegetation Science.* 5:703-714.

Robinson, D.C.E. 1998. Prognosis Southern Interior (SI 1.0) Geographic Variant Description (Draft). Prepared by ESSA Technologies, Vancouver BC for BC Ministry of Forests, Victoria, BC. 162 pp.

Utzig, G. 1997. Terrain and Soil Inventory: West Arm Demonstration Forest (WADF). Unpublished Report for Kootenay Lake Forest District, BC Ministry of Forests, Nelson, BC. 33 pp. with Maps and App.

Wykoff, W.R., N.L. Crookston, and A.R. Stage. 1982. User's guide to the Stand Prognosis Model. USDA Forest Service, Intermountain Forest and Range Experiment Station, Ogden, UT. Gen. Tech. Rep. INT-133. 112 pp.