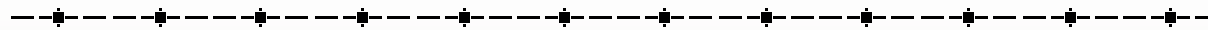
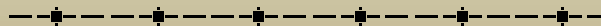


REGENERATION PREDICTION FOR PROGNOSIS^{BC}, IDFdm2 SUBZONE VARIANT, INVERMERE FOREST DISTRICT



Katrina Froese, M.Sc. (candidate), Valerie LeMay, Ph.D., RPF,
Peter Marshall, Ph.D., RPF, and Abdel-Azim Zumrawi, Ph.D., RPF

Forest Resources Management Department
University of British Columbia



MY WORK

✦ PRIMARY OBJECTIVES:

- ✦ Create regeneration imputation models for the IDFdm2 study area
- ✦ Compare performance of tabular, MSN, and k-MSN approaches

✦ OUT OF CURIOSITY:

- ✦ Compare imputation results to predictions using the original FVS model, NI variant

✦ LOOKING FOR INPUT:

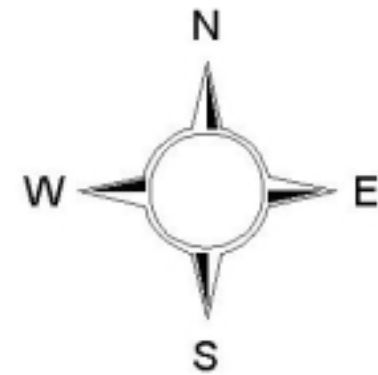
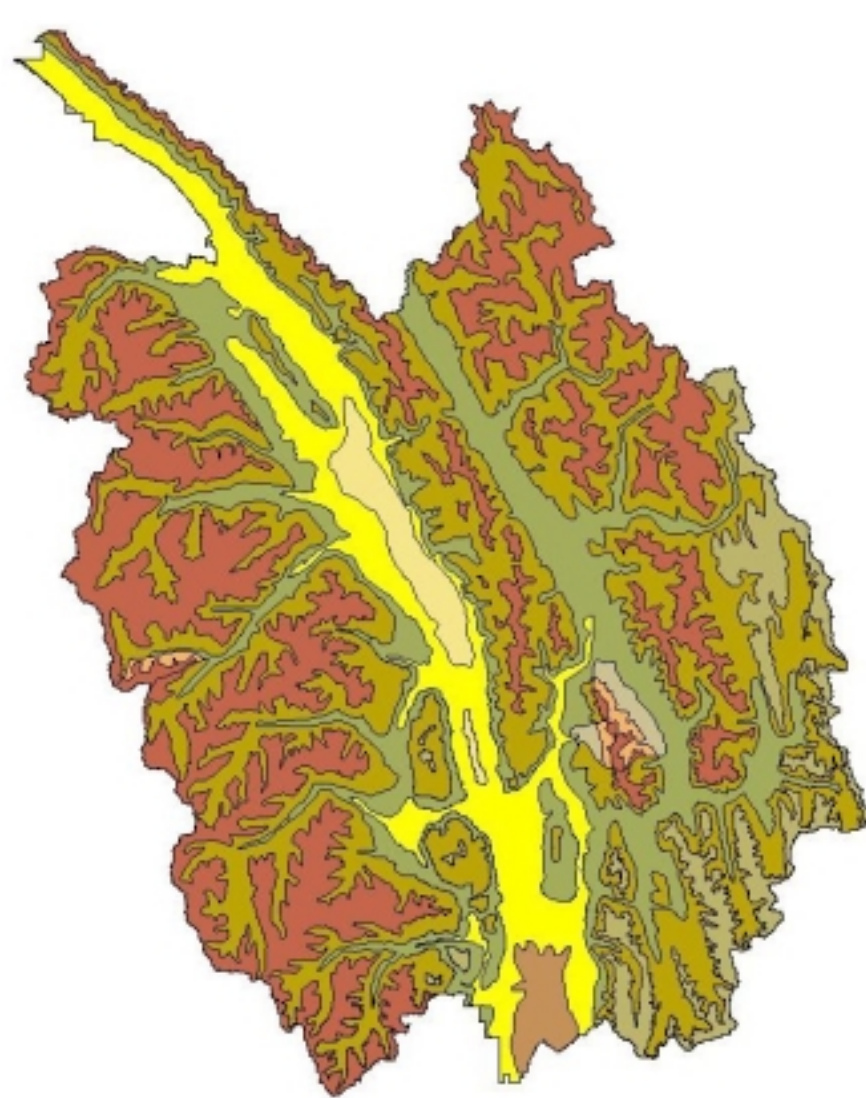
- ✦ Relating spatially mapped data to regeneration clumps

BACKGROUND

-
- ✦ Recalibration of FVS NI for the ICHmw2, Nelson Forest Region, resulted in improved estimates but results were limited by small sample sizes
 - ✦ Recent efforts have focused on imputation techniques as an alternative method for predicting regeneration

STUDY AREA

- ✦ IDFdm2, Invermere Forest District, Nelson Forest Region, southeastern British Columbia
- ✦ Valley bottoms and lower slopes, Rocky Mountain Trench, 900-1200m elevation
- ✦ Stands are Douglas-fir (*Pseudotsuga menziesii* var. *glauca*) or Douglas-fir with:
 - ◆ ponderosa pine (*Pinus ponderosa*)
 - ◆ lodgepole pine (*Pinus contorta* var. *latifolia*)
 - ◆ western larch (*Larix occidentalis*)
 - ◆ hybrid white spruce (*Picea glauca* x *Engelmannii*)
 - ◆ trembling aspen (*Populus tremuloides*)
 - ◆ paper birch (*Betula papyrifera*)
- ✦ Generally uneven aged stands, managed with partial cutting (some clearcuts)



Bec.shp

-  AT
-  ESSFdk
-  ESSFdkp
-  ESSFwm
-  ICHmk1
-  IDFdm2
-  IDFun
-  MSdk
-  PPdh

25 0 25 50 Kilometers

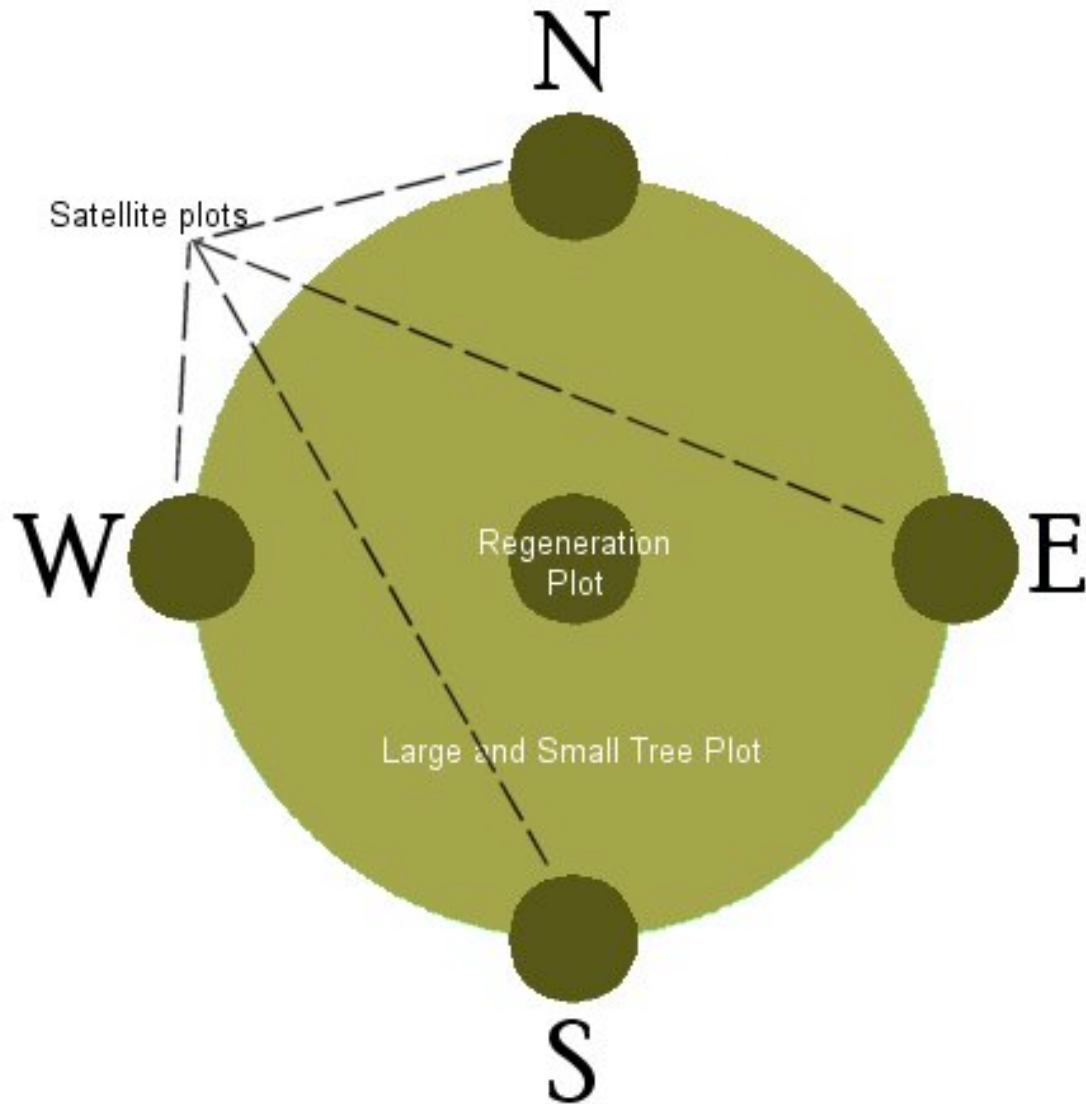


Location of biogeoclimatic variants within the Invermere Forest District

SAMPLING FRAME

- ✦ All areas within the IDFdm2 study area disturbed within the last 5-25 years
- ✦ Candidate openings selected from sampling matrix based on:
 - ◆ Number of years since disturbance
 - ◆ Silvicultural system
 - ◆ BEC site series
 - ◆ Elevation
- ✦ Good geographic range of sites was obtained
 - ◆ North to south along trench
 - ◆ East to west

SAMPLING DESIGN





DISTRIBUTION AND NUMBER OF PLOTS

-
- ✦ A total of 111 plots from 37 polygons were sampled
 - ✦ 89 from partially cut stands
 - ✦ 11 from clearcuts
 - ✦ 11 from undisturbed stands

IMPUTATION: METHODS

- ✦ Five regeneration sub-plots were averaged (n=111 plots)
- ✦ For all models, variables of interest to be imputed were regeneration sph by species and height class
 - Class 1: 15-49.95 cm (shade tolerant & semi-tolerant species)
 - Class 2: 49.95-99.95 cm
 - Class 3: 99.95-129.95 cm
 - Class 4: >129.95 cm and <7.5 cm DBH
- ✦ Data split into 5 data sets with 80% reference and 20% target/test – 5 runs to assess performance
- ✦ Final tabular models and final fit statistics were based on the full data set
- ✦ Any tabular imputation tables created based on n=1 plots were discarded due to lack of reliability

MEASURES OF MODEL PERFORMANCE

-
- ✦ Mean Absolute Deviation
 - ✦ Bias (positive bias=underprediction, negative bias=overprediction)
 - ✦ Root Mean Squared Error
 - ✦ Presence: proportion of plots with presence correctly predicted
 - ✦ Absence: proportion of plots with absence correctly predicted

TABULAR IMPUTATION: MODEL 1

- ✦ Three moisture classes
 - ◆ dry, mesic, wet
- ✦ Two basal area classes
 - ◆ open $<20\text{m}^2/\text{ha}$, dense $\geq 20\text{m}^2/\text{ha}$
- ✦ Five years since disturbance classes
 - ◆ 5-9 years, 10-14 years, 15-19 years, 20-24 years, undisturbed
- ✦ Planted vs. unplanted
- ✦ 10 categories with $n=1$, 9 categories with $n=0$

TABULAR IMPUTATION: MODEL 1

- ✦ Average regeneration per hectare, unplanted sites, for years since disturbance class 3 (15-19 years), basal area class “open” and moisture class “mesic”.

Moisture Class	Number of Plots	Species	Height Class				Total
			1	2	3	4	
mesic	12	At		0	0	0	0
		Bl	0	0	0	0	0
		Ep		50	12	0	62
		Fd	3145	1895	693	1870	7603
		Lw		12	0	50	62
		Pl		37	25	173	235
		Py		74	0	37	111
		Sxw	12	12	0	12	37
		Total		3158	2080	731	2142

TABULAR IMPUTATION: MODEL 2

- ✦ Two basal area classes
 - ◆ open $<20\text{m}^2/\text{ha}$, dense $\geq 20\text{m}^2/\text{ha}$
- ✦ Five years since disturbance classes
 - ◆ 5-9 years, 10-14 years, 15-19 years, 20-24 years, undisturbed
- ✦ Planted vs. unplanted
- ✦ No moisture classes
- ✦ No categories with $n=1$ or $n=0$

TABULAR IMPUTATION: MODEL 2

- ✱ Average regeneration per hectare, unplanted sites, for years since disturbance class 3 (15-19 years), and basal area class “open”.

Number of Plots	Species	Height Class				Total
		1	2	3	4	
17	At		74	25	74	173
	Bl	0	0	0	0	0
	Ep		111	161	211	483
	Fd	2229	1263	483	1238	5213
	Lw		12	0	50	62
	Pl		25	37	173	235
	Py		0	0	25	25
	Sxw	12	12	0	12	37
	Total	2241	1498	706	1783	6229

MSN and k-MSN IMPUTATION

- ✦ Imputation done using MSN 2.0 software (Crookston *et al.* 2002)
- ✦ Weighted Euclidean distances based on canonical correlations between the X and Y variables (developed by Moeur and Stage 1995)
- ✦ Run separately for each species
- ✦ For MSN, predicted regeneration sph is from the neighbour with the shortest distance
- ✦ For k-MSN, predicted regeneration sph is from the average of the three most similar neighbours

MSN and k-MSN IMPUTATION

X VARIABLE	DESCRIPTION
EASTING	<i>UTM easting (m)</i>
NORTHING	<i>UTM northing (m)</i>
COSASPECT	<i>cosine of aspect(in radians) x percent slope</i>
SINASPECT	<i>sine of aspect(in radians) x percent slope</i>
SL	<i>percent slope</i>
ELEV	<i>elevation (m)</i>
ELEVSQ	<i>elevation squared</i>
MOIST	<i>moisture class (dry, mesic, wet)</i>
TPH	<i>trees per hectare</i>
BAHA	<i>basal area of stand (m²)</i>
CCF	<i>crown competition factor</i>
OM	<i>presence of organic material (Y/N)</i>
PLANT	<i>is site planted (Y/N)</i>
INT	<i>intermediate entries such as thinning or pruning (Y/N)</i>
PREP	<i>site preparation following harvesting (Y/N)</i>
YRSINCE	<i>number of years since harvest</i>

RESULTS: FIT STATISTICS

Root Mean Squared Error (sph)

	Tabular1	Tabular2	MSN	k-MSN
At	482	542	675	555
Bl	37	43	79	57
Ep	262	357	262	297
Fd	4066	4063	5793	4747
Lw	182	189	181	178
Pl	1169	1175	852	801
Py	81	80	80	85
Sxw	100	300	260	275
Total	3918	3977	5897	4521

RESULTS: FIT STATISTICS

Mean Absolute Deviation (sph)

	Tabular1	Tabular2	MSN	k-MSN
At	198	257	234	192
Bl	10	12	17	14
Ep	78	129	70	75
Fd	2943	2928	4325	3457
Lw	74	81	75	65
Pl	602	622	439	386
Py	29	30	19	25
Sxw	57	137	92	96
Total	2942	2903	4399	3431

RESULTS: FIT STATISTICS

- ✦ Dominant species: highest MAD, RMSE and bias
- ✦ Sporadic species: low MAD, RMSE and bias
- ✦ Intermediate species: moderate values
- ✦ Overall, Tabular Model 2 fit statistics were poorer than Tabular Model 1; however, it outperformed Tabular Model 1 during data splitting runs
- ✦ Fit statistics for k-MSN consistently outperformed those for MSN

RESULTS: MATCHING ACCURACY

	Presence			
	Tabular1	Tabular2	MSN	k-MSN
At	0.36	0.24	0.37	0.33
Bl	0.15	0.10	0.00	0.00
Ep	0.30	0.19	0.33	0.40
Fd	0.89	0.89	0.86	0.89
Lw	0.25	0.20	0.11	0.14
Pl	0.61	0.53	0.47	0.49
Py	0.24	0.19	0.30	0.28
Sxw	0.33	0.25	0.34	0.35
Total	0.97	0.97	0.98	0.97

RESULTS: MATCHING ACCURACY

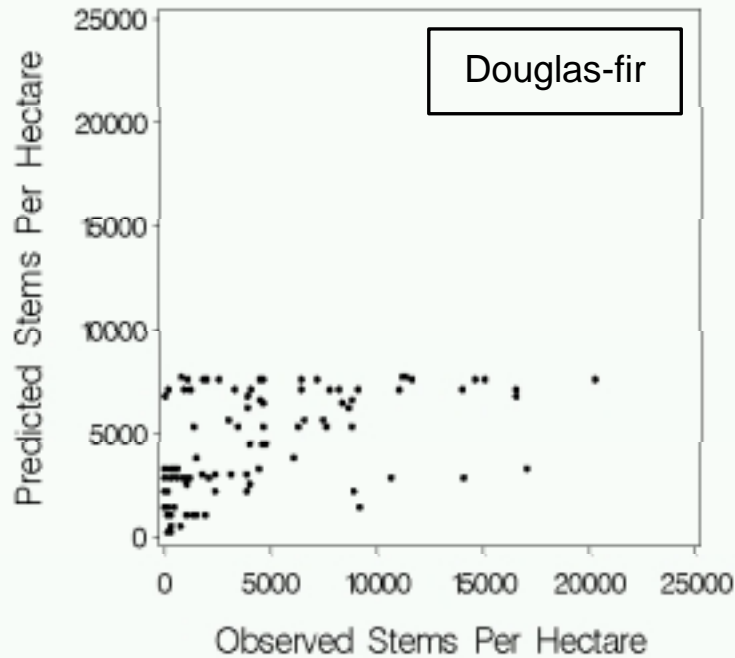
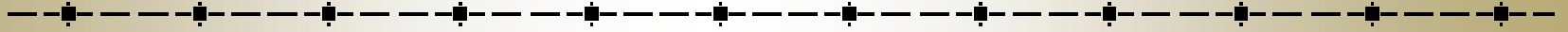
Absence

	Tabular1	Tabular2	MSN	k-MSN
At	0.53	0.22	0.81	0.70
Bl	0.83	0.76	0.94	0.91
Ep	0.83	0.71	0.94	0.94
Fd	0.00	0.00	0.17	0.00
Lw	0.52	0.32	0.78	0.64
Pl	0.45	0.28	0.53	0.34
Py	0.72	0.66	0.94	0.88
Sxw	0.49	0.16	0.79	0.67
Total	0.00	0.00	0.33	0.00

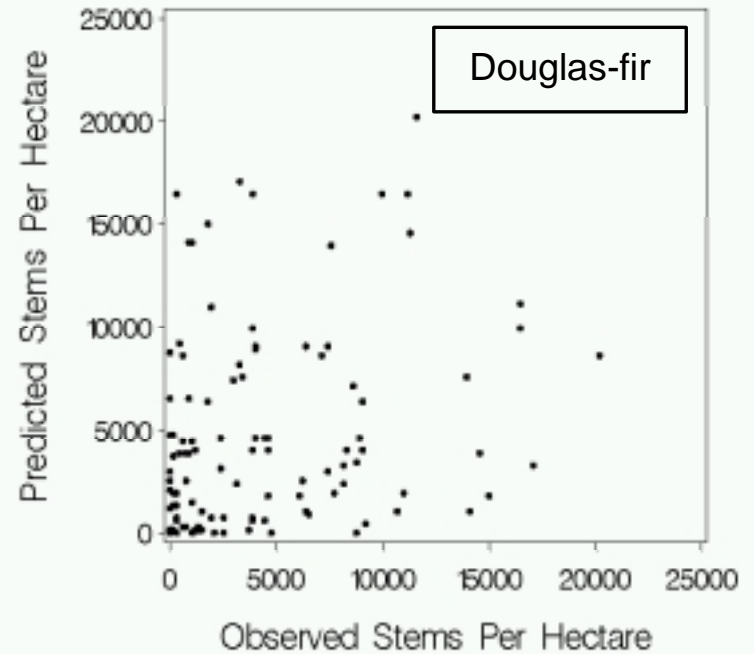
RESULTS: PREDICTION ACCURACY

- ✦ Dominant species: good prediction of presence and poor prediction of absence
- ✦ Sporadic species: poor prediction of presence and good prediction of absence
- ✦ Intermediate species: moderate prediction of both presence and absence
- ✦ Tabular Model 2 generally lower prediction accuracy for presence and absence than Model 1
- ✦ Predicting presence was similar for k-MSN and MSN, but MSN was better in predicting absence

RETENTION OF VARIABILITY



Tabular Model 1



MSN

MY THOUGHTS

- ✦ Imputation models are easy to use and update
- ✦ BUT small sample sizes can still mean poor prediction
- ✦ Intended use should guide which method is preferred (Tabular vs. MSN or k-MSN)
- ✦ Larger sample sizes would allow more classes for tabular imputation – more variability retained
- ✦ Try selecting variables for tabular imputation classes based on canonical correlations

FVS vs. IMPUTATION: METHODS

- ✦ Regeneration model (Ferguson and Carlson 1993) programmed in SAS 8.02
- ✦ Site Series 01 & drier = Habitat Type 1 (PSME/VAGL)
- ✦ Site Series 01/04 & wetter = Habitat Type 5 (ABGR/LIBO)
- ✦ Predicted sph by species only (no heights/classes)
- ✦ No addition of planted trees (Fd, Pl, Lw)

RESULTS: PREDICTION ACCURACY

	Root Mean Squared Error (sph)		
	FVS1	Tabular2	k-MSN
At	614	542	555
Bl	93	43	57
Ep	387	357	297
Fd	5188	4063	4747
Lw	430	189	178
Pl	1259	1175	801
Py	283	80	85
Sxw	341	300	275
Total	5070	3977	4521

RESULTS / MY THOUGHTS

- ✦ FVS NI overpredicts both the presence and abundance of Lw, Py and PI
- ✦ Large bias (underprediction) of Fd - may relate to different minimum heights
- ✦ Provides no predictions for hardwoods (At, Ep) but fit statistics are similar
- ✦ Calibration of the original FVS regen model still has the potential to provide good predictions
- ✦ FVS still requires more labour intensive sampling per plot, and is difficult to model/update

SPATIALLY MAPPED DATA: SAMPLING METHODS

-
- ✦ One in four 11.28 m radius plots were spatially mapped (for a total of 24 plots)
 - ✦ Distance and bearing to:
 - ◆ Large trees (DBH, species)
 - ◆ Small trees (DBH, species)
 - ◆ Regeneration clumps (species, height class, length, width, axis)
 - ◆ Stumps (diameter at 0.15 m)
 - ◆ Windthrow (DBH, length, axis)
 - ◆ Large slash piles (as regen)
 - ◆ Large clumps of shrubs (as regen)
 - ✦ Plus substrate information collected on regen subplot

IDEAS FOR THE DATA

- ✦ Competitive indices
- ✦ Clumpiness indices
- ✦ Measures of protection for regeneration.
 - ◆ Possibly “area of protection available” – size of opening around regen clumps, position within opening?
- ✦ Issues:
 - ◆ Fixed area plots – limits the radius around clumps
 - ◆ Small sample size
 - ◆ So what? How useful is knowing that there is a spatial relationship?

MY THOUGHTS

-
- ✦ If spatial information is important to modelling, we may have to rethink how we want to allocate our data collection and modelling efforts
 - ✦ If we can show that there is a relationship to clumpiness factors, there may be a role for imputing clumpiness factors from aerial data rather than requiring inventories

We gratefully acknowledge the following partners and collaborators:

Barry Phillips, Research Branch; Barry Snowden and Ralph Winter, Forest Practices Branch; Jon Vivian, Inventory Branch; Cliff Beliveau, Darrell Regimbald, Richard Dominy, and the staff of the Invermere Forest District; Grant Neville and Douglas Braybrook of Tembec Industries; Vivian Jablanczy of Slocan Forest Products; and Tim Harding of Riverside Forest Products.

Funding for this research provided by Forest Renewal BC and Forestry Innovation Investment. In-kind contributions were supplied by the BC Ministry of Forests, the University of BC Faculty of Forestry, Tembec Industries, Slocan Forest Products, and Riverside Forest Products.

REFERENCES:

Crookston, N.L., M. Moeur, and D. Renner. 2002. Users guide to the most similar neighbour imputation program, Version 2. USDA For. Serv. RMRS-GTR-96. 35 pp.

Ferguson, D.E. and C.E. Carlson. 1993. Predicting regeneration establishment with the Prognosis Model. USDA Forest Service. Intermountain Research Station, Ogden, Utah, Research Paper INT-467. 54 pp.

Moeur, M. and A.R. Stage. 1995. Most similar neighbour: an improved sampling inference procedure for natural resource planning. For. Sci. 41(2): 337-359.

Robinson, D.C.E. and W.A. Kurtz. 1998. Analysis of the Prognosis regeneration establishment model behaviour with BC regeneration data. Prepared by ESSA Technologies Ltd., Vancouver, BC for the BC Ministry of Forests, Victoria, BC. 42 pp.