Forest Ecology and Management 177 (2003) 277–285

Forest Ecology
and
Management

www.elsevier.com/locate/foreco

Imputing tree-lists from aerial attributes for complex stands of south-eastern British Columbia

By H. Temesgen, V.M. LeMay*, K.L. Froese, P.L. Marshall

Department of Forest Resources Management, University of British Columbia,
2045-2424 Main Mall, Vancouver, BC, Canada V6T 1Z4

Received 2 April 2002; accepted 17 June 2002

Abstract

The nearest neighbor, *k*-nearest neighbors, distance-weighted *k*-nearest neighbor, and class-weighted *k*-nearest neighbor imputation methods were compared for accuracy in estimating tree-lists (list of species and diameter for each tree) from aerial attributes for complex stands, with up to nine species and a wide range of sizes, in south-eastern British Columbia, Canada. For the four imputation methods, the most similar neighbor distance metric was used, and three neighbors were used for the *k*-nearest neighbor methods. Ground variables used to represent the tree-list included the number of trees per hectare by species, ranges of diameter by species, and basal area per hectare. Aerial variables included species composition, crown closure (%), elevation, biogeoclimatic ecosystem classification (BEC) zones, height, age, and site class. Sample data were divided, and the imputation methods were compared for accuracy using observed and estimated species composition, stand tables, basal area, and volume per hectare. Also, the imputed tree-list was used to predict yield using a stand level growth model, and this predicted yield was compared to the yield obtained using the actual tree-list. Of the four approaches used, the nearest neighbor was marginally better, but the methods that averaged the three nearest neighbors were somewhat better for the distribution of stems per hectare by diameter for the more sparse hardwood species. Of the three averaging methods, weighting by similarity of the species composition and the BEC zone provided better results. In using the estimated trees lists in a growth and yield model, the average volumes were reasonable at the beginning and end of the period for all methods. However, the volumes for a particular stand could be quite different than that obtained for an observed tree-list.

© 2002 Elsevier Science B.V. All rights reserved.

Keywords: Imputation; Tree-lists; Aerial data; Complex stands

1. Introduction

Increasingly, resource management requires information linked across stand and forest (landscape) levels to assess sustainability, and to obtain information needed for certification. To realize this linkage, analysts would like detailed information on every stand within a management unit, including a detailed tree-list (list of species and diameter for every tree). However, this is often not possible because of the cost and time needed to obtain detailed information within stands.

In British Columbia (BC), forests cover 59 million hectares and most of this forest land is mountainous with few roads. Due to the large forested land base and difficult access, tree-lists are available only for the limited number of ground-sampled stands. However,

---

[*] Corresponding author. Tel.: +1-604-822-4770; fax: +1-604-822-9106.
E-mail address: lemay@interchange.ubc.ca (V.M. LeMay).

0378-1127/02/$ – see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0378-1127(02)00321-3

stand level aerial information such as crown closure, height class and species composition, is available for every stand. This information may be used to estimate tree-lists where they are not available.

Tree-lists can be used as inputs to tree and stand projection models to update inventories, develop landscape and forest management plans, and to assess stand structure at the stand level. Several tree and stand projection models (growth and yield models) have been developed for use in BC. Some of these growth and yield models (e.g. Prognosis^{BC}, adapted from the original model developed by Stage (1973) and the mixed-wood growth model (MGM) developed by Titus at the University of Alberta, Edmonton) require a tree-list to set initial conditions for the stand being modeled or projected. Tree-lists can also be used as inputs to habitat models (e.g. Ohmann et al., 1994; Beukema et al., 2000). Tree-lists for tree and stand projection models may only include live trees, whereas tree-lists for habitat models may focus on live and dead trees, and on particular size groupings. Dead standing and recently dead and fallen trees (i.e. wildlife trees) contribute to forest structure, dynamics and forest succession (Bull et al., 1997; Ganey, 1999) and provide habitat for the maintenance or enhancement of wildlife (e.g. nest cavities, nurseries, etc.) (Ohmann et al., 1994; Ganey, 1999).

Most research on generating a tree-list from stand attributes have used stand level information measured on the ground, and have been used on simple stands with few species and uni-modal diameter distributions. The approaches used can be broadly categorized into diameter distribution (frequency by species and diameter) modeling and imputation methods.

The diameter distribution modeling approach involves fitting a diameter distribution for each stand, and then predicting the parameters of the distribution using stand variables. Depending upon how the parameters are predicted, the diameter distribution modeling approaches in the literature have been classified as parameter prediction (Rennolls et al., 1985; Biging et al., 1994, Lindsay et al., 1996; Von Gadow and Hui, 1999), parameter recovery (Bailey and Dell, 1973; Bailey, 1980; Hyink and Moser, 1983), and percentile prediction (Bailey, 1980; Maltamo et al., 2000). These approaches are based on uni-modal distributions, and do not accurately describe the diameter distributions of complex stands that have multi-modal and irregular diameter distributions.

Imputation is defined as "replacing missing or non-sampled measurements for any unit in the population with measurements from another unit with similar characteristics" (Van Deusen, 1997). The process for generating tree-list information is to locate the closest neighbor (most similar stand level variables) for the stand that does not have the tree-list information (called target stands), from a pool of stands that have detailed tree and stand data (reference stands). Imputation methods include the most similar neighbor (Moeur and Stage, 1995; Moeur, 2000), the k-nearest neighbor (Maltamo and Kangas, 1998), geo-statistical estimation (Moeur and Hershey, 1998), and tabular imputation methods (Ek et al., 1997). Unlike the diameter distribution modeling approach, imputation methods can retain both spatial and attribute variance structures of the data (Ek et al., 1997; Moeur and Stage, 1995), do not restrict the form or shape of the underlying distribution, and will always result in projections within the bounds of biological reality (Moeur and Stage, 1995). This approach is therefore more applicable for complex stands, than diameter distribution modeling.

In this paper, four imputation methods were used to estimate tree-lists from stand level aerial data for stands of south-eastern BC. Although aerial data and tree-lists from ground data were available for every stand in the study, the data were divided into reference stands (aerial data and tree-list available) and target stands (aerial data only) to compare the four methods. The ground data that were available for the target stands were used to assess the accuracy of imputation. For this study, the tree-list included the live trees only; however, these approaches could be used for other definitions of tree-lists more suitable for habitat models.

2. Methods

2.1. Data

Ground and aerial data for 96 complex stands collected in 1996 for south-eastern BC were used for this study (Table 1). The stands included several tree species: Douglas-fir (*Pseudotsuga menziesii* (Beissn.) Franco), lodgepole pine (*Pinus contorta* var. *contorta* Dougl.), western white pine (*Pinus monticola* Dougl.), Ponderosa pine (*Pinus ponderosa*

Table 1
Distribution of the 96 sampled stands by crown closure, height, age, and biogeoclimatic ecosystem classification zone[a]

Crown closure (%)	20	30	40	50	60	70	80
Number of stands	5	19	14	23	21	12	2
Height (m)	15	25	35				
Number of stands	32	48	16				
Age (years)	70	90	110	130	180		
Number of stands	24	24	12	9	27		
Ecological (BEC) zone	AT	ESSF	ICH	IDF	MS		
Number of stands	3	26	52	7	8		

[a] AT: the alpine tundra zone, ESSF is the engelmann spruce/subalpine-fir zone; ICH: the interior cedar-hemlock zone; IDF: the interior Douglas-fir zone, and MS is montane spruce zone.

Laws.), whitebark pine (*Pinus albicaulis* Englem.), western larch (*Larix occidentalis* Nutt., L), trembling aspen (*Populus tremuloides* Michx.), subalpine-fir (*Abies lasiocarpa* (Hook.) Nutt.), spruce (*Picea glauca* (Moench) Voss and *P. engelmannii* Parry and hybrids), black cottonwood (*Populus trichocarpa* Torr. & Gray), western hemlock (*Tsuga heterophylla* (Raf.) Sarg.), white birch (*Betula paperifera* Marsh.), and western red cedar (*Thuja plicata* Donn.). The aerial variables available from inventory databases for each stand were percent crown closure, percent composition by species, height class, age class, and site index. Height classes 2–4 were assigned heights of 15, 25, and 35 m, respectively, and age classes 4–8 were assigned ages of 70, 90, 110, 130 and 180 years, respectively. Using the aerial information (i.e. height, crown closure, site index and species composition), a stand level growth and yield model, variable density yield projection model (VDYP) (BC, 1995), was used to obtain estimates of volume per hectare, average height, and quadratic mean diameter for each stand. For the ground data, four variable radius plots were randomly located in each stand, and the species, diameter outside bark at breast height (1.3 m above ground; DBH), and status (i.e. live or dead) for all trees with a DBH of 12.5 cm or greater were recorded, along with other tree variables (BC, 1998). The live trees for all ground data were compiled, and for each stand, the average tree-list (i.e. list of stems per hectare by species and DBH) was calculated. In addition, the average volume per hectare, stems per hectare, and basal area per hectare were obtained for all species combined, and by species.

The 96 sampled stands were randomly divided into reference (74; 77% of sampled stands) and target (22; 23% of sampled stands) datasets. A routine in SAS, Version 6.12, PROC IML, was written and used to obtain the matching stand from the reference list for each target stand based on the MSN program developed by Moeur et al. (1999). A set of 22 proxy variables were used to represent the tree-list (*Y* variables) and related to a set of 14 aerial variables (*X* set) (Table 2). The measure of similarity used was similar to the Mahalanobis distance, except that it was weighted using the results of canonical correlation analysis between the ground (*Y*) and aerial variables (*X*), as outlined in Moeur and Stage (1995):

$$D_{uj}^2 = (X_u - X_j)\Gamma\Lambda^2\Gamma'(X_u - X_j)'$$

Table 2
Ground and aerial variables used in the imputations[a]

Ground variables (*Y* set)	Aerial variables (*X* set)
Stems per hectare for seven species: Douglas-fir, cedar, lodgepole pine, spruce, larch, subalpine-fir, and hardwoods[b]	Percent composition for seven species
	Height (m)
	Age (years)
	Site index (m)
Minimum and maximum diameter (cm) for each of seven species	Crown closure (%)
	Average height (m), model estimated
	Quadratic mean diameter (cm), model estimated
Basal area per hectare for all species (m^2/ha)	Volume (m^3/ha), model estimated

[a] Ground variables were selected as a proxy for the tree-list.

[b] Hardwoods included aspen, cottonwood, and birch.

where Γ is the matrix of standardized canonical coefficients for the aerial variables; Λ^2 the diagonal matrix of squared canonical correlations between aerial attributes and ground variables; X_u a vector of standardized values of the aerial variables for the uth target stand; and X_j a vector of standardized values of aerial variables for the jth reference stand. The use of the $\Gamma\Lambda^2\Gamma'$ matrix results in weightings of the aerial variables, depending on the relationship with the ground variables. A smaller value for the distance metric indicates that the variables are more similar. This most similar neighbor (MSN) distance metric was selected over the absolute (Maltamo and Kangas, 1998) and Euclidean distances (Moeur, 2000), as it considers the relationships between aerial and ground inventory variables in weighting the distances.

Although only one measure of similarity was used, four different ways of weighting the measure of similarity were compared.

1. The nearest neighbor, that is, the neighbor having the lowest distance (NN). All stands in the reference list were ranked by similarity. The tree-list for the most similar stand was used as the imputed tree-list for the target stand.
2. Un-weighted average of three reference stands (k-NN). The tree-lists of the three most similar reference stands were selected and averaged to represent the tree-list of the target stand, as was done in Maltamo and Kangas (1998). They showed how the number of selected nearest neighbor stands used in finding the average influenced the accuracy of an imputation method for pure and mixed Scots pine (*Pinus sylvestris*) stands. For this study with up to nine species in a stand, three stands were selected for the averaging. More than three stands would have resulted in possibly very dissimilar stands being used in the averaging. Less than three stands might have resulted in species distributions that were not as similar to the target stand.
3. Distance-weighted average of three reference stands (NN_WT). The tree-lists of the three most similar reference stands were used to calculate a weighted average tree-list, using the inverse of the similarity as the weight.
4. Class-weighted average of three reference stands (NN_CWT). The tree-lists of the three most similar reference stands were used to calculate a weighted average tree-list. The weight used in this case was the similarity of the species composition and the BEC zone between the target and reference stands, rather than the inverse of the similarity measure. A weight of: (1) three was used if both the reference and target stands had a similar species composition and the same BEC zones; (2) two was used if the reference and target stands have either the same species composition or the same BEC zones, not both; and (3) one was used if the reference and target stand had neither the same species composition nor the same BEC zones. Hardin (1994) used a similar weighting approach in assigning land classes to remotely sensed data.

2.2. Comparison of approaches

The accuracy of the four approaches for estimating the tree-lists for the 22 target stands was measured using bias and root mean square error (RMSE) of the each of the ground variables used in the imputation (Table 1), as was done by Maltamo and Kangas (1998). The bias and RMSE values for each variable were obtained by:

$$\text{Bias} = \frac{\sum_{i=1}^{n}(\text{observed}_i - \text{imputed}_i)}{n}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(\text{observed}_i - \text{imputed}_i)^2}{n}}$$

where n is the number of target stands. Biases in stems per hectare were also calculated by DBH class for all species, and separately for coniferous versus deciduous species. Graphical comparisons of the observed (target) and estimated stems per hectare by DBH class and species were also used to assess the imputation results.

Since one goal of generating a tree-list is to project the stand growth and yield, the estimated tree-lists from each of the four approaches were used as input into the Prognosis$^{\text{BC}}$ model to project volume over time, from the time of measurement (1996) to 50 years in the future (2046), and these projections were compared to the projection using the actual tree-list. The average, minimum, and maximum volumes at 1996 and at 2046 were compared for the four methods versus using the actual tree-list. Three of the stands were classified as alpine tundra (AT) in the dataset. However, the AT zone generally has few trees and, therefore, these were set to montane spruce (MS) for the Prognosis^{BC} runs, based on the ground sample data.

3. Results and discussion

The 96 sampled stands covered a wide range for the ground-measured variables, including 12.5–217.0 cm DBH for live trees, and from 166 to 2414 stems per hectare. The aerial variables ranged from 7 to 34 m in site index, 20–80% in crown closure, and 16.3–42.9 cm in predicted quadratic mean diameter from the VDYP model.

Since there were 14 aerial variables and 22 ground variables, 14 canonical covariates resulted from the canonical correlation analysis. Canonical correlations were very strong with 0.944, 0.90, and 0.85 for the first, second, and third canonical variates, respectively. The first six variates had canonical correlations >0.60. Quadratic mean diameter had high coefficients for five of these six coefficients. Other variables that weighted highly were projected height, and model estimated volume. The percent composition of Douglas-fir, lodgepole pine, hardwoods, larch, and true fir had high coefficients on a few covariates. Other variables such as site, age, and crown closure had lower coefficients on the first six covariates, and therefore, were weighted less in the Most Similar Neighbor distance measurement.

Unlike the results reported by Maltamo and Kangas (1998) for more simple Scots pine stands, NN resulted in lower biases (absolute values) in the stems per hectare by species, than the other methods, except for hemlock and larch, and also lower biases in basal area per hectare and volume per hectare (Table 3). Generally, NN resulted in lower RMSE values also. The species composition by stems per hectare appeared to be better estimated using NN. For the other methods using the average of three stands as the inputted tree-list, the differences in weights did not change the bias of any of the variables substantially. An unexpected result was that the use of NN_CWT did not improve the estimated species compositions. The better results using NN might have been due to the large sample of stands available for imputation of the 22 target stands. A suitable match was found in the stand list. Using simulations of repeated sampling on the same dataset, LeMay and Temesgen (2001) noted that an increase from 20 to 50% of reference stands resulted in improvements in estimating, whereas an increase from 50 to 80% produced little improvements for NN. This sampling intensity might not be possible with most forest inventories; improvements in the

Table 3
Bias and root mean squared error (RMSE) of four imputation approaches for a selection of stand level ground variables for live tree-lists[a]

Ground variable	Bias				RMSE			
	NN	k-NN	NN_WT	NN_CWT	NN	k-NN	NN_WT	NN_CWT
Stems per hectare for								
Douglas-fir	−83	146	146	147	232	225	229	223
Lodgepole pine	48	204	204	205	185	264	260	267
Spruce	48	81	81	81	123	107	107	109
Cedar	36	68	68	69	91	93	93	93
Hemlock	−32	−1	−2	0	58	50	50	50
Larch	−38	2	1	3	4	4	41	40
Hardwoods	7	26	25	25	4	4	31	32
All species	−14	527	523	532	302	624	618	631
Basal area per hectare (m^2/ha)	0	24	24	24	17.1	28	28	28
Volume per hectare (m^3/ha)	26	251	249	253	92.9	279	278	280

[a] Imputation methods are: the closest neighbor (NN), average of the three most similar neighbors (k-NN), average of the three most similar neighbors weighted by similarity (inverse distance; NN_WT), and average of the three most similar neighbors weighted by similarity of BEC zone and species group (NN_CWT) ($n = 22$ stands).

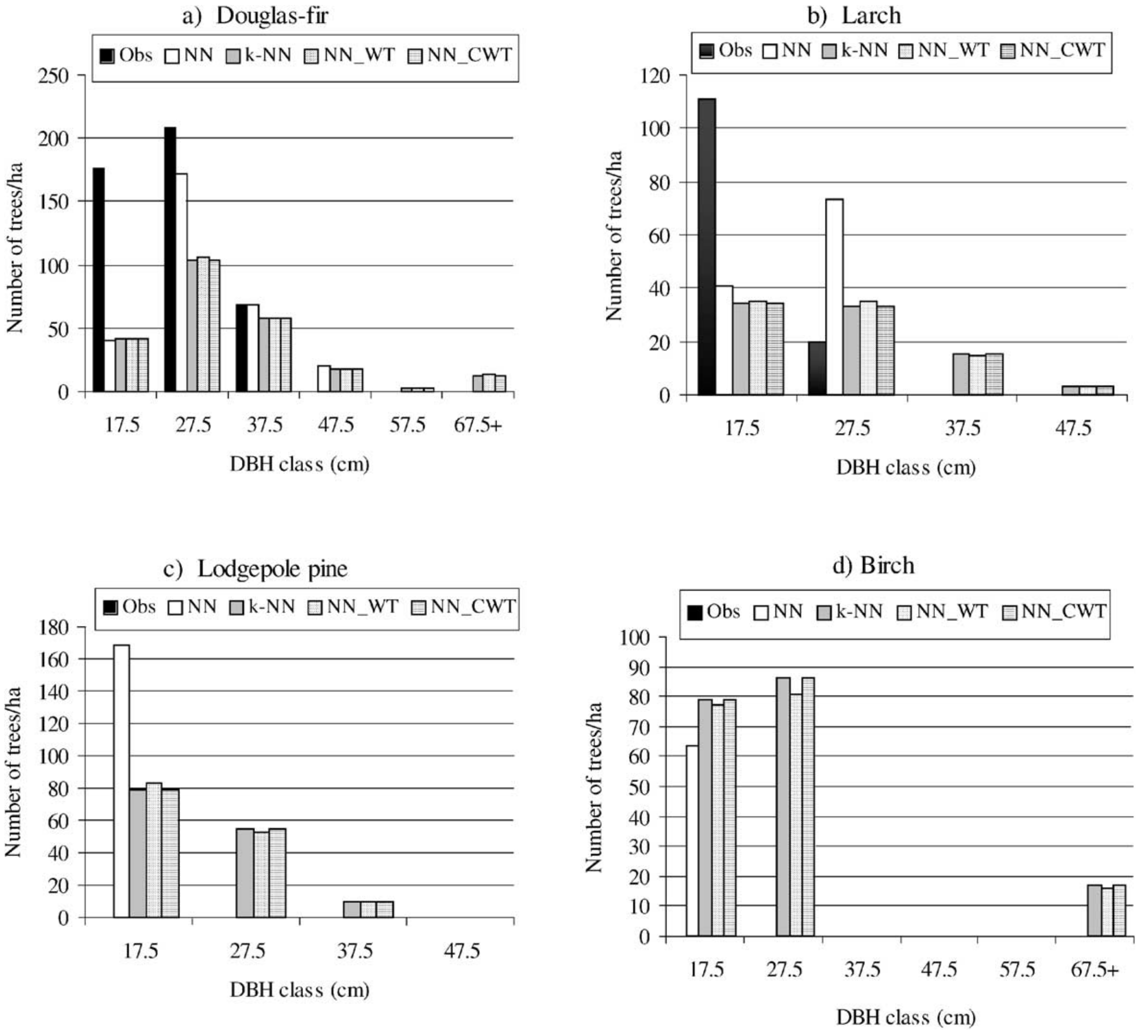

Fig. 1. Observed (black) and estimated diameter frequencies by species for stand 345. NN (white), *k*-NN (grey), NN_WT (spotted), and NN_CWT (striped) are the most similar neighbor stand, averages of three reference stands, and weighted averages of three reference stands by inverse distance, and by matches on species composition and ecological (BEC) zone, respectively.

species composition of inputted tree-lists might be obtained through the posterior use of the aerial species proportions to alter the results. A similar approach was used by Puumalainen et al. (2001) to alter predicted DBH distribution, using auxiliary variables with good results.

For estimated stems per hectare by DBH class, there is no clear advantage for any of the four methods tested for all species combined (Fig. 1 and Table 4). For coniferous species only, the NN resulted in lower biases, and lower maximum differences (absolute values; Table 4). Of the methods using the average of three stands, the NN_CWT method gave better results, by weighting using the species composition and BEC zone. Deciduous species comprised much smaller proportion of the stand than coniferous species; the averaging methods gave better results than NN for these rare species (Table 4). Biases in all three cases tended to be higher for the smallest DBH class (Table 4).

Using the inputted tree-lists as inputs to Prognosis^{BC}, the volume per hectare over the projection period

Table 4
Bias and minimum (min), and maximum (max) differences in estimated stems per hectare by DBH class for each of the four imputation methods, for (a) all species; (b) coniferous species only; and (c) hardwood species only[a]

DBH class	NN			*k*-NN			NN_WT			NN_CWT		
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
(a) All species												
17.5	36	−700	536	22	−258	677	24	−266	672	21	−317	683
27.5	1	−299	248	9	−173	414	9	−199	406	9	−150	422
37.5	−18	−197	155	3	−85	200	3	−108	211	3	−72	190
47.5	3	−52	60	4	−46	116	3	−46	116	4	−46	116
57.5	−11	−50	39	6	−34	61	5	−36	59	6	−32	64
67.5	6	6	6	11	−23	98	11	−23	98	11	−22	98
(b) Coniferous species												
17.5	29	−700	536	21	−258	677	22	−266	672	19	−317	683
27.5	−2	−299	248	11	−173	414	11	−199	406	11	−150	422
37.5	−18	−197	155	4	−85	200	4	−108	211	4	−72	190
47.5	3	−52	60	4	−46	116	4	−46	116	4	−46	116
57.5	−11	−50	39	6	−34	61	5	−36	59	6	−32	64
67.5	6	6	6	12	−23	98	12	−23	98	12	−22	98
All	−32	−763	1286	67	−1052	2414	68	−1080	2414	122	−629	856
(c) Hardwood species												
17.5	51	−113	283	45	−61	283	45	−61	283	45	−62	283
27.5	5	−79	70	−10	−121	70	−9	−113	70	−11	−128	70
37.5	−18	−41	−8	−7	−14	−3	−8	−16	−3	−7	−14	−2
47.5	−28	−28	−28	−9	−9	−9	−10	−10	−10	−9	−9	−9
57.5	−7	−24	9	1	−8	9	1	−8	9	1	−8	9
67.5	−14	−14	−14	−7	−7	−7	−7	−7	−7	−7	−7	−7
All	7	−323	283	26	−45	303	25	−26	139	25	−26	138

[a] Imputation methods are: the closest similar neighbor (NN), average of the three most similar neighbors (*k*-NN), average of the three most similar neighbors weighted by similarity (inverse distance; NN_WT), and average of the three most similar neighbors weighted by similarity of BEC zone and species group (NN_CWT) ($n = 22$ stands).

was examined. Trends varied by stand, and results were not satisfactory for every stand. For example, the volumes using the inputted tree-lists were about 100 m^3/ha more than that using the actual tree-list for stand 186 at the beginning of the projection period (Fig. 2). However, reasonable results were obtained, on average, with the average volume for the target stands of 382 m^3/ha for 1996 using the actual tree-list (observed), and the average estimated volumes from 335.1 to 363.4 m^3/ha, depending upon the method used in imputation (Table 5). Averages were also similar for the end of the projection period (2046). The standard deviation of volumes was lower using the imputation methods relative to the observed tree-lists, for the 1996 and 2046 estimates. Likely, these lower standard deviations in the estimated volumes were due to one stand with a volume of 1040.5 m^3/ha using the observed tree-list in 1996 (maximum volume for observed). All of the imputation methods had a lower maximum volume. For all imputation methods, an unusual stand in the target stands that is not represented in the reference list results in a poorly estimated tree-list for that stand. This also explains the lower average volumes using the imputed tree-lists, but this average volume was not substantially different from the average volume using the observed tree-lists. On average, the imputation methods provided good results, even if a particular stand was not well estimated.

Overall, the NN approach appeared to give better results than other methods based on averaging three stands, particularly for species comprising a high proportion of the stand composition. For rarer species, averaging using a match on the species proportion and BEC zone (NN_CWT) gave slightly better results for stems per hectare by DBH class.

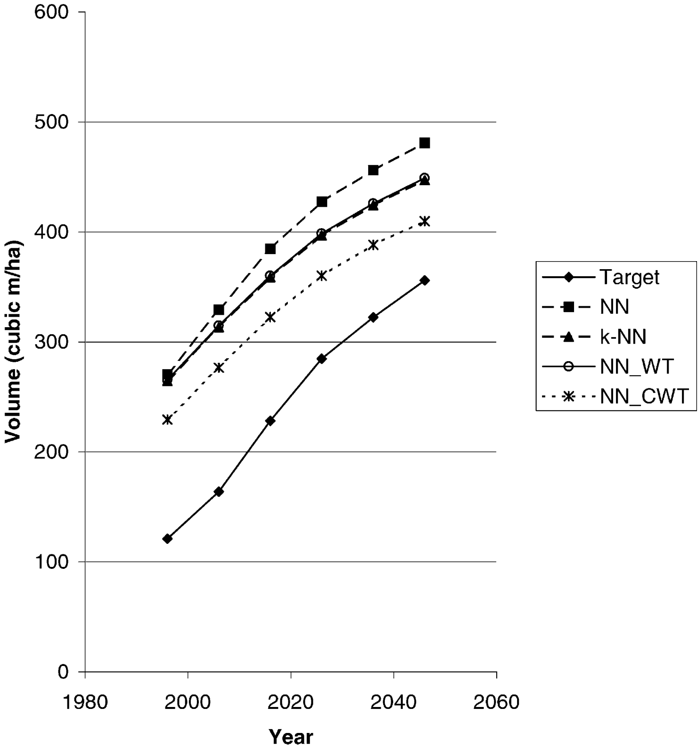

Fig. 2. Volume over stand age based on Prognosis^{BC} projections of observed (target) and imputed tree-lists for stand 186. NN, *k*-NN, NN_WT, and NN_CWT are the most similar neighbor stand, averages of three reference stands, and weighted averages of three reference stands by inverse distance, and by matches on species composition and ecological (BEC) zone, respectively.

Table 5
Mean, minimum (min), standard deviation (S.D.), and maximum (max) of estimated volumes (m^3/ha) using the observed tree-list, versus using the estimated tree-list from each of the four imputation methods. Imputation methods are: most similar neighbor (NN), average of the three most similar neighbors (*k*-NN), average of the three most similar neighbors weighted by similarity (inverse distance; NN_WT), and average of the three most similar neighbors weighted by similarity of BEC zone and species group (NN_CWT) ($n = 22$ stands)

Year/method	1996				2046			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
Observed	382.0	198.2	175.3	1040.5	496.4	176.6	285.1	1004.9
NN	335.2	90.4	194.9	481.1	447.3	114.9	239.7	663.4
k-NN	358.6	130.9	194.9	784.4	452.7	114.0	215.9	674.6
NN_WT	363.4	119.8	124.8	612.5	482.2	103.3	285.1	636.8
NN_CWT	340.1	91.8	198.6	482.5	452.2	115.0	243.1	667.6

4. Conclusions

Tree-list information on every parcel of land provides better flexibility for forest management systems. For very complex stands with multiple species and a wide variety of tree sizes, imputation methods may better estimate the tree-list than other estimation methods, such as diameter distribution modeling, that are more suited to uni-modal distributions of stems over DBH classes.

Of the four approaches used, NN was marginally better, but the methods that average the three nearest neighbors were somewhat better for the distribution of stems per hectare by DBH for the more rare hardwood species. Of the three averaging methods, weighting by similarity of the species composition and the BEC zone provided better results. In using the estimated tree-lists in a growth and yield model, the average volumes over all stands were reasonable at the beginning and end of the period. However, the volumes for a particular stand could be quite different than that obtained for an observed tree-list.

Acknowledgements

We gratefully acknowledge the cooperation and support provided by the Ministry of Forests, Forest Inventory and Forest Practices Branches. This project was funded by Forest Renewal BC and administered by BC Ministry of Forests, Forest Practices Branch.

References

Bailey, R.L., 1980. Individual tree growth derived from diameter distribution model. For. Sci. 26, 626–632.

Bailey, R.L., Dell, T.R., 1973. Quantifying diameter distribution with the Weibull function. For. Sci. 19, 97–104.

BC Ministry of Forests. 1995. VDYP Interactive Application User Guide, Version 6.3. Victoria, BC, Canada, 39 pp.

BC Ministry of Forests, 1998. Inventory Audit Sampling Standards and Procedures. Victoria, BC, Canada, 40 pp.

Beukema, S.J., Reinhardt, E., Greenough, J.A., Kurz, W.A., Crookston, N., Robinson, D.C.E., 2000. Fire and fuels extension: model description. Internal Report for USDA. Forest Service, Rocky Mountain Research Station, Moscow, ID.

Biging, G.S., Robards, T.A., Turnblom, E.C., VanDeusen, P.C., 1994. The predictive models and procedures used in the forest stand generator. STAG Hilgardia 61 (1), 1–36.

Bull, E.L., Parks, C.G., Torgersen, T.R., 1997. Trees and logs important to wildlife in the interior Columbia River Basin. USDA Forest Service General Technical Report, PNW-GTR-391. Pacific Northwest Research Station, Portland, OR.

Ek, A.R., Robinson, A.P., Radtke, P.J., Walters, D.K., 1997. Development and testing of regeneration imputation models for forests in Minnesota. For. Ecol. Manage. 94, 129–140.

Ganey, J.L., 1999. Snag density and composition of snag populations on two national forests in northern Arizona. For. Ecol. Manage. 117, 169–178.

Hardin, P.J., 1994. Parametric and nearest-neighbor methods for hybrid classification: a comparison of pixel assignment accuracy. Photog. Eng. Rem. Sens. 60, 1439–1448.

Hyink, D.M., Moser Jr., J.W., 1983. A generalized framework for projecting forest yield and stand structure using diameter distributions. For. Sci. 29, 85–95.

LeMay, V., Temesgen, H., 2001. Nearest neighbor methods for generating tree-lists from aerial data. In: Proceedings of the Invited Paper Presented at the IUFRO 4.11 Forestry Biometry, Modeling, and Information Science Conference. The University of Greenwich, School of Computing and Mathematical Sciences, Greenwich, England, 26–29 June 2001, 23 pp. (abstract posted at http://cms1.gre.ac.uk/conferences/infro/).

Lindsay, S.R., Wood, G.R., Woollons, R.C., 1996. Stand table modeling through the Weibull distribution and usage of skewness information. For. Ecol. Manage. 81, 19–23.

Maltamo, M., Kangas, A., 1998. Methods based on *k*-nearest neighbor regression in the prediction of basal area diameter distribution. Can. J. For. Res. 28, 1107–1115.

Maltamo, M., Kangas, A., Uuttera, J., Torniainen, T., Saramaki, J., 2000. Comparison of percentile based prediction and Weibull distribution in describing the diameter distribution of heterogeneous Scots pine stands. For. Ecol. Manage. 133, 263–274.

Moeur, M., 2000. Extending stand exam data with most similar neighbor inference. In: Proceedings of the Society of American Foresters National Convention, Portland, OR, 11–15 September 1999, pp. 99–107.

Moeur M., Hershey, R., 1998. Preserving spatial and attribute correlation in the interpolation of forest inventory data. In: International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Quebec City, Que., Canada, 20–22 May 1998, pp. 419–430.

Moeur, M., Stage, A.R., 1995. Most similar neighbor: an improved sampling inference procedure for natural resource planning. For. Sci. 41, 337–359.

Moeur M., Crookston, N., Renner, D., 1999. Most Similar Neighbor, Release 1.0. User's Manual. USDA Forest Service, Rock Mountain Research Station, Moscow, ID, 48 pp.

Ohmann, J.L., McComb, W.C., Zumrawi, A., 1994. Snag abundance for primary cavity-nesting birds on non-federal forestlands in Oregon and Washington. Wildl. Soc. Bull. 22, 607–620.

Puumalainen, J., Maltamo, M., Kangas, A., 2001. Calibrating predicted diameter distribution with additional information for structurally diverse forest stands. In: Proceedings of the Continuous Cover Forestry: Assessment, Analysis, and Scenarios IUFRO Conference, Göttingen, Germany, 19–21 September 2001, pp. 47–58.

Rennolls, K., Geary, D.N., Rollinson, T.J.D., 1985. Characterizing diameter distributions by the use of Weibull distribution. Forestry 58, 57–66.

Stage, A.R., 1973. Prognosis model for stand development. USDA, Forest Service Report Paper INT-137. Ogden, UT, 32 pp.

Van Deusen, P.C., 1997. Annual forest inventory statistical concepts with emphasis on multiple imputations. Can. J. For. Res. 27, 379–384.

Von Gadow and G. Hui 1999. Modeling Forest Development. Kluwer Acadamic Publisher, Dordrecht, The Netherlands, 213 pp.