EVALUATION OF REGION AND SUBREGION-BASED TOTAL VOLUME EQUATIONS FOR DAHURIAN LARCH IN NORTHEAST CHINA

MBANGILWA, M. M.[1*] – JIANG, L. C.[1*]

[1]Key Laboratory of Sustainable Forest Ecosystem Management-Ministry of Education, School of Forestry, Northeast Forestry University, Harbin 150040, China

*Corresponding authors
e-mail: jlichun@nefu.edu.cn

Abstract. A total of 10 sub-regions were identified and tree total volume equations were developed for economically important Dahurian larch species in northeast China (Daxing'an Mountain region/Heilongjiang province). These equations were based on a simple power function and adjusted to the forest inventory data volumes of three mountainous regions, including the northwest of northern slope of Yilehuli Mountains (NWYLHLM), the southeast of northern slope of Yilehuli Mountains (SAYLHLM) and the eastern slope of northern part of Daxin'an Mountains (NDXAM). The volume equations were tested to determine if differences by ecoregion were statistically significant. Results varied by species, volume and ecoregion suggesting that the relationship between tree volume and tree diameter of Dahurian larch (*Larix gmelinii* Rupr) tree species is significantly different between the three geographic regions, depending on local climatic, soil and ecological conditions. This implies that the provincial and regional models of volume-diameter relationships are not appropriate for predicting tree volume at ecoregional level. In addition, the use of an ecoregional specific model will lead to significant bias in predicting the total volume of trees. The ecoregion-based volume models developed in this study could provide more accurate information on tree growth and development to forest resource managers and planners.
Keywords: *Larix gmelinii, Heilongjiang, forest inventory, mountainous regions, ecoregions*

Introduction

Boreal forests, the second largest biome in the world, cover about one-third of the Earth's forest area (Jia and Zhou, 2018). Dahurian larch (*Larix gmelinii* Rupr) is a dominant tree species in Chinese boreal forest. It's one of the main timber species in China and forms large forests in northeast China, including the northeastern part of the Inner Mongolia autonomous region and the Daxing'an Mountain range (Jin et al., 2015). It's the most commercially cultivated timber species in northeastern China due to its ecological prevalence and its superior wood attribute. However, its timber quality is largely driven by the crown architecture, i.e., the number, size and distribution of branches. The majority of branch-level models in the literature are focused on planted forests, which have substantially different crown architecture than that grown in natural mixed forests (Dong et al., 2016). Therefore, the purpose of this study is to develop the total volume model for Dahurian larch in three forest regions of northeastern China.

Accurate estimation and determination of stem standing volume is very useful for both research and practical purposes in forestry and contributes to the efficient and sustainable management of timber resources (Barrio et al., 2007). When preparing a cut for sale, the estimate of the volume of trees to be felled is necessary. The cubage can be done on a tree, as in the case of the volume of large trees and can also be done on a stand with an estimate of the overall volume of wood made by type of wood. The cubing of standing timber is governed by the standard Afnor NF B 53-017 "Estimated

volume of standing timber" for estimating the "solid volume" of a standing tree (Limousin, 2011). Previous studies have shown that, total volume equation has been widely used throughout eastern North America (Honer, 1967). In a study conducted in 1998 (Shailer et al., 1998) in New Brunswick, Honer's volume equations were shown to best predict total volume from a set of almost 20 different published volume equations for nine important commercial species.

Recently Lee et al. (2017) reported that, Stem volume models using diameter at breast height (DBH) and height were developed for *Pinus densiflora*, *Pinus koraiensis*, and *Larix kaempferi*. The data were obtained from Gangwon, Gyeonggi, and North Gyeongsang provinces in South Korea, and the sample trees felled were used to provide the parameter estimates of volume equations. The combined-variable function, $V = a + bD^2H$, was shown to be the best model through the validation of the equation. Also, the model using only DBH, $V = aD + bD^2$ was also evaluated to be applicable in the field. Similarly, Tsega et al. (2018) reported as well that, Stem volume equations were fitted for *Cupressus lusitanica* in Gergeda Forest, Ethiopia using six different established equation forms. A total of 260 trees were measured for their diameter at breast height (*D*), total height (*H*) and stem volume using destructive sampling methods. The data set was randomly divided into equal size for equation development and equation validation. Five fit statistics comprising of the fit index, root mean square error, bias (%), absolute mean deviation and coefficient of variation were used to evaluate the performance of each equation. Among the different equations, the Schumacher and Hall function of the form $V = b_1 D^{b_2} H^{b_3}$, which estimates volume (*V*) using diameter at breast height and total height as predicting variables, performed best and was then fitted to the combined data set for prediction of volume over-bark of *C. lusitanica* in Gergeda Forest.

These models revealed higher accuracy when compared with previous studies. Both equations are considered to be easily used in the field. In the recent past, the principles of multifunction, multipurpose and ecologically based forest management were approved in China. Therefore, in order to help forest managers and practitioners realize the differences in these relationships among different sub-regions then subregion-based height-diameter models are very important.

Previous studies have also revealed that, there is ecoregion differences exist in the height-diameter relationships in Ontario (Zhang et al., 2002; Peng et al., 2004) and in West Virginia (Brooks and Wiant, 2008; Jiang et al., 2005) and local merchantable board foot, cubic foot and weight equations have been developed for six economically important hardwood species in central Appalachia. Equations were based on a simple power function and fit to volumes from the USDA Forest Service Forest Inventory Analysis data for West Virginia and parts of Ohio, Pennsylvania, Maryland, Virginia, and Kentucky. Five ecoregions were identified and the volume equations were tested to determine if differences by ecoregion were statistically significant with results varying by species, volume type and ecoregion (Brooks and Wiant, Jr., 2004). Akinsanmi and Akindélé (1995) noted that, sustainable management of forest resources requires adequate data collection. These actions make it possible for the manager to have indicators on wood potential in order to develop a consequent management plan. Palm (1982), Pauwels and Rondeux (1999) and Thibaut et al. (1998), revealed that the low values of the statistical parameters of validation are obtained with the formula of the truncated cone. The graphical analysis of the residues of the truncated cone model shows a good distribution of the residues. Consequently, Fonton et al. (2009) confirmed

the hypothesis of the succession of cone trunks along the stem of *Isoberlinia spp, Anogeissus leiocarpa* and *Danielli aoliveri* stems from the study by on the modeling of drum volume tree for sustainable management of forest ecosystems. Rondeux (1999), Laumans (1991), Thibaut et al. (1998) and Fonton et al. (2002) agreed with the findings of the models obtained in the Fonton study, expressing the volume as a function of diameter squared (D^2), are the most relevant for estimating individual tree volume. Finally, Bontemps et al. (2012) found that the estimation of forest-impaired biomass and forest productivity explained that environmental changes involve redefining and expanding the ecosystem services provided by forests, while aiming for the sustainability of their management. The quantification of timber and wood energy resources and the carbon storage function requires new methods of assessing forest volume and biomass.

In this current study, the test is based on the ecoregional statistical analysis from forest management data in Dahurian larch region, in the Daxing'an mountainous region located in northeastern China (Heilongjiang Province). For the purpose of this study again, there are many situations in forest mensuration when the use of local volume tables is useful (Brooks and Wiant, 2008). These may include but are not limited to cursory forest inventory applications as well as timber trespass evaluations when stem diameter is estimated from an existing stump diameter. This study uses forest inventory data from the Daxing'an Mountains region in Heilongjiang Province in northeastern China.

The objectives of this study are to evaluate whether ecoregion based volume cubic meter, total volume equations are statistically justified to test for differences between ecoregions for the one species selected and to evaluate the bias associated when a region wide model is compared to individual ecoregion models. Consequences of inappropriately applying provincial, regional or other regulations were also evaluated among the ten subregions.

Materials and Methods

Study area and data

The study was conducted in the cold temperate forest regions in Heilongjiang Province (Daxing'an Mountains; from 121° 12′E to 127° 00′E and from 50° 10′N to 53° 33′N) and eastern part of Inner Mongolia, in Northeastern China) (*Figure 1*). The elevation of the area ranges from 300 to 700 m above the sea level. The soil in this region is mostly Haplum-brepts or Eutroboralfs (or dark brown forest soil in Chinese Taxonomic System). The area experinces continental monsoon climate with mean annual rainfall ranges from 500 to 750 mm and mean annual temperature is from -1 to -2.8°C.

These temperate forests are dominated by White birch (*Betula platyphylla*), Amur linden (*Tilia amurensis*), Maple (*Acer mono*), Dahurian birch (*Betula davuria*), Mongolianoak (*Quercus mongolica*), Dahurian poplar (*Populus davidiana*), and mixed hardwood forest dominated by Amurlinden (*Tilia amurensis*), Maple (*Acer mono*), Manchurian ash (*Fraxinus manshurica*), Manchurian walnut (*Juglans mandshurica*), Mongolian oak (*Quercus mongolica*) and Manchurian elm (*Ulmus laciniata*) (Dong et al., 2014).

Data and methods

Ten subregions for conifer trees were identify for this study based on the work conducted by Brooks and Wiant (2008) in USA (*Table 1*).

Table 1. *F-test for testing the difference between subregions*

Subregions	n	Full Model			Reduced Model			Extra sum of squares		
		SSE_F	PF	df_F	SSE_R	PR	df_R	$df_R - df_F$	F-value	P-value
Overal	2411	26.2484	30	2384	32.3012	3	2408	24	22.9060	0.0000
SL - JGDQ	738	8.0276	6	732	8.2117	3	735	3	5.5972	0.0008
SL - XL	404	5.6357	6	398	7.0271	3	401	3	32.753	0.0000
SL - TH	417	3.0916	6	413	3.3481	3	414	3	11.370	0.0000
SL - HZ	429	4.0404	6	423	5.5989	3	426	3	54.389	0.0000
SL - SBZ	451	3.4801	6	445	3.4992	3	448	3	0.8120	0.4876
SL - HJY	342	2.9661	6	336	3.1289	3	339	3	6.1497	0.0004
SL - XLJ	438	2.4011	6	432	2.5310	3	435	3	7.7918	0.0000
SL - TQ	405	5.2095	6	399	5.9319	3	402	3	18.442	0.0000
SL - AME	499	8.2247	6	493	8.3661	3	496	3	2.8245	0.0382
JGDQ - XL	714	9.4562	6	708	11.218	3	711	3	43.966	0.0000
JGDQ - TH	727	6.9120	6	721	7.1381	3	724	3	7.8624	0.0000
JGDQ - HZ	739	7.8608	6	733	9.5449	3	736	3	52.343	0.0000
JGDQ - SBZ	761	7.3006	6	755	7.4159	3	758	3	3.9758	0.0079
JGDQ - HJY	652	6.7865	6	646	6.8625	3	649	3	2.4089	0.0660
JGDQ - XLJ	748	6.2216	6	742	6.5218	3	745	3	11.933	0.0000
JGDQ - TQ	715	9.0299	6	709	9.9681	3	712	3	24.553	0.0000
JGDQ - AME	809	12.045	6	803	12.426	3	806	3	8.4645	0.0000
XL - TH	393	4.5201	6	387	4.7256	3	390	3	5.8629	0.0006
XL - HZ	405	5.4689	6	399	5.8731	3	402	3	9.8294	0.0000
XL - SBZ	427	4.9087	6	421	6.5642	3	424	3	47.329	0.0000
XL - HJY	318	4.3947	6	312	4.7316	3	315	3	7.9746	0.0000
XL - XLJ	414	3.8297	6	408	4.9515	3	411	3	39.839	0.0000
XL - TQ	381	6.6381	6	375	6.8315	3	378	3	3.6418	0.0129
XL - AME	475	9.6533	6	469	11.221	3	472	3	25.394	0.0000
TH - HZ	418	2.9248	6	412	3.6358	3	415	3	33.384	0.0000
TH - SBZ	440	2.3646	6	434	2.6741	3	437	3	18.936	0.0000
TH - HJY	331	1.8505	6	325	1.8929	3	328	3	2.4795	0.0611
TH - XLJ	427	1.2855	6	421	1.7295	3	424	3	48.458	0.0000
TH - TQ	394	4.0939	6	388	4.1496	3	391	3	1.7599	0.1543
TH - AME	488	7.1099	6	482	7.3279	3	485	3	4.9428	0.0021
HZ - SBZ	452	3.3131	6	446	5.3358	3	449	3	90.744	0.0000
HZ - HJY	343	2.7993	6	337	3.5644	3	340	3	30.703	0.0000
HZ - XLJ	439	2.2343	6	433	3.6992	3	436	3	94.628	0.0000
HZ - TQ	406	5.0427	6	400	5.8446	3	403	3	21.202	0.0000
HZ - AME	500	8.0579	6	494	9.6820	3	497	3	33.187	0.0000
SBZ - HJY	365	2.2391	6	359	2.4153	3	362	3	9.4182	0.0000
SBZ - XLJ	461	1.6741	6	455	1.8099	3	458	3	12.303	0.0000
SBZ - TQ	428	4.4825	6	422	5.1044	3	425	3	19.517	0.0000
SBZ - AME	522	7.4977	6	516	7.6256	3	519	3	2.9323	0.0330
HJY - XLJ	352	1.1600	6	346	1.5273	3	349	3	36.511	0.0000
HJY - TQ	319	3.9684	6	313	4.1198	3	316	3	3.9795	0.0083
HJY - AME	413	6.9837	6	407	7.1959	3	410	3	4.1227	0.0067
XLJ - TQ	415	3.4035	6	409	4.0321	3	412	3	25.182	0.0000
XLJ - AME	509	6.4187	6	503	6.6461	3	506	3	5.9405	0.0005
TQ - AME	476	9.2271	6	470	10.009	3	473	3	13.277	0.0000

Note: N – sample size, SSE_F – error sum of squares of the full model, df_F – degrees of freedom of SSE_F, SSE_R – error sum of squares of reduced model, df_R – degrees of freedom of SSE_R, P is the number of parameters.

A total of 5474 individual height-diameter measurements for four Conifer Forest tree species were collected in systematically/randomly selected square/circular (25.8 m x 25.8 m plots) using field Inventory from Daxing'an Mountains in the northeast of China (*Figure 1*). A county was assigned as part of a specific ecoregion if a majority of the land base fell within ecoregion boundaries. The selected ecoregions are a combination of the different subregions between them and the land base regions identified by the China Soil Conservation Service. In this study, only one species of tree, the Dahurian larch (*Larix gmelinii* Rupr*)* was selected, given the absence of some data for other tree species.

Figure 1. The geographical location of study area and plot distribution in the Northeast China

A total of 2411 destructively sampled Dahurian larch trees were used in this investigation. These trees were felled throughout the forest inventory areas of northeast China and all sampled trees were selected to ensure a representative distribution across a range of height and diameter classes within stands varying in density, height, site condition, age and stand structure. The data was then randomly divided into an adaptation dataset and a validation dataset (*Table 2*). For developing taper and individual tree volume equations the data set was initially used. Diameters at breast height were measured for all sampled trees (D, defined as 1.3 m above the ground) outside bark and total volume (V). Trees possessing broken tops, obvious cankers or crooked boles were excluded from the analysis. Summary statistics for tree diameter and total volume are provided for each subregion, the NDXAM, YLHLM-NW, and YLHLM-SA regions, and all data combined (Overall) in *Table 3*.

The three major regions employed are depicted in *Figure 1* above and include:

- Region1: The northwest of the northern slope of Yilehuli Mountains (NWYLHLM) which includes four subregions: Xilinji (XLJ), Tuqiang (TQ), Amuer (AME), Huzhong (HZ).
- Region2: The southeast of the northern slope of Yilehuli Mountains (SAYLHLM) which includes four subregions: Xinlin (XL), Tahe (TH), Shibazhan (SBZ), Hanjiayuan (HJY).

- Region3: The eastern slope of northern part of Daxing'an Mountains (NDXAM), which includes two subregions: Songling (SL), Jiagedaqi (JGDQ).

Table 2. *Summary of fitting and validation data sets for Dahurian larch (Larix gmelinii Rupr.) for the ten subregions in Northeastern China*

Subregions	Statistic	Fitting data			Validation data		
		D (cm)	Height (m)	Volume (m³)	D (cm)	Height (m)	Volume (m³)
SL	N°of observations	214	214	214	214	214	214
	Mean	29.15	19.59	0.81	25.87	40.21	0.70
	Minimum	5.50	7.10	0.01	4.10	7.50	0.01
	Maximum	55.10	29.20	2.55	50.20	87.50	2.28
	SD	14.07	5.41	0.70	12.88	19.56	0.62
JGDQ	N°of observations	524	524	524	524	524	524
	Mean	27.27	18.30	0.66	23.99	37.55	0.54
	Minimum	5.40	6.40	0.01	4.50	7.00	0.01
	Maximum	61.00	30.80	2.90	59.40	83.20	2.71
	SD	13.24	4.83	0.60	12.01	17.66	0.52
XL	N°of observations	190	190	190	190	190	190
	Mean	30.94	19.34	0.87	27.95	42.22	0.76
	Minimum	6.00	8.40	0.01	4.90	9.00	0.01
	Maximum	63.40	29.50	3.48	58.60	89.00	3.20
	SD	12.57	3.99	0.69	11.69	17.10	0.62
TH	N°of observations	203	203	203	203	203	203
	Mean	21.98	15.34	0.45	20.34	28.80	0.40
	Minimum	5.20	5.10	0.01	4.20	7.20	0.00
	Maximum	50.60	26.50	2.22	48.00	64.00	2.13
	SD	12.32	5.37	0.52	11.83	15.12	0.47
HZ	N°of observations	215	215	215	215	215	215
	Mean	23.53	16.13	0.50	20.37	31.11	0.41
	Minimum	5.10	4.50	0.01	4.10	6.00	0.01
	Maximum	48.80	25.90	2.34	45.80	64.50	2.04
	SD	11.61	4.69	0.51	10.07	14.73	0.67
SBZ	N°of observations	237	237	237	237	237	237
	Mean	27.46	17.93	0.64	24.49	39.13	0.54
	Minimum	5.40	6.10	0.01	4.40	6.40	0.01
	Maximum	56.60	26.10	2.11	50.60	91.00	1.84
	SD	13.33	4.72	0.54	12.30	19.50	0.47
HJY	N°of observations	128	128	128	128	128	128
	Mean	26.79	17.72	0.62	23.78	37.11	0.52
	Minimum	5.20	6.60	0.01	3.90	7.50	0.01
	Maximum	50.80	26.50	2.37	46.80	78.20	2.08
	SD	12.38	4.60	0.55	11.40	17.94	0.47
XLJ	N°of observations	224	224	224	224	224	224
	Mean	16.00	15.76	0.24	14.24	22.66	0.20
	Minimum	5.10	7.20	0.01	3.80	6.50	0.01
	Maximum	62.00	24.60	2.15	46.00	82.00	1.60
	SD	9.00	4.21	0.29	8.03	13.81	0.25
TQ	N°of observations	191	191	191	191	191	191
	Mean	23.98	17.45	0.62	21.86	33.34	0.54
	Minimum	5.20	6.50	0.01	4.60	6.40	0.01
	Maximum	56.00	28.50	2.70	53.00	96.00	2.42
	SD	14.59	5.25	0.68	13.74	20.63	0.60
AME	N°of observations	285	285	285	285	285	285
	Mean	33.68	22.37	1.08	30.36	47.93	0.95
	Minimum	5.10	7.33	0.01	4.60	6.30	0.01
	Maximum	72.20	33.98	3.99	62.60	89.10	3.39
	SD	13.00	4.87	0.75	11.70	18.70	0.67

D = diameter at breast height, SD=Standard of deviation

The individual tree database for each region was queried to select only those trees from the appropriate counties as well as selecting only those merchantable tree records having measured diameters and heights and non-zero volumes. The tree data selected was from the northwest of the northern slope of Yilehuli Mountains, the southeast of the northen slope of Yilehuli Mountains, the eastern slope of northern part of Daxin'an Mountains, located in Heilongjiang Province. The volume parameters examined included only total volume in cubic meter, as reported in the tree database. The number of trees species and ecoregions sampled after the combination of different ecoregions is shown in *Table 1*, and the state distribution parameter and the summary Tree diameter (DBH) statistics within volume (V) for regional and ecoregional datasets for Dahurian larch species are presented in *Table 3*. An initial evaluation was conducted using several families of nonlinear models to establish an appropriate volume model form (*Table 4*).

Table 3. *Summary statistics of tree diameter at breast (DBH), total height (H) and total volume (V) for regional and subregional data sets*

Ecoregions	DBH (cm)					H (m)				V (m³)			
	N	Mean	STD	Min	Max	Mean	STD	Min	Max	Mean	STD	Min	Max
SL	214	29.15	14.07	5.50	55.10	19.59	5.41	7.10	29.20	0.81	0.70	0.01	2.55
JGDQ	524	27.27	13.24	5.40	61.00	18.30	4.83	6.40	30.80	0.66	0.60	0.01	2.90
XL	190	30.94	12.57	6.00	63.40	19.34	3.99	8.40	29.50	0.87	0.69	0.01	3.48
TH	203	21.98	12.32	5.20	50.60	15.34	5.37	5.10	26.50	0.45	0.52	0.01	2.22
HZ	215	23.53	11.61	5.10	48.80	16.13	4.69	4.50	25.90	0.50	0.51	0.01	2.34
SBZ	237	27.46	13.33	5.40	56.60	17.93	4.72	6.10	26.10	0.64	0.54	0.01	2.11
HJY	128	26.79	12.38	5.20	50.80	17.72	4.60	6.60	26.50	0.62	0.55	0.01	2.37
XLJ	224	16.00	9.00	5.10	62.00	15.76	4.21	7.20	24.60	0.24	0.29	0.01	2.15
TQ	191	23.98	14.59	5.20	56.00	17.45	5.25	6.50	28.50	0.62	0.68	0.01	2.70
AME	285	33.68	13.00	5.10	72.20	22.37	4.87	7.33	33.98	1.08	0.75	0.01	3.99
NWYLHLM	**915**	**24.94**	**13.86**	**5.10**	**72.20**	**18.26**	**5.53**	**4.50**	**33.98**	**0.64**	**0.68**	**0.01**	**3.99**
SAYLHLM	**758**	**26.75**	**13.10**	**5.20**	**63.40**	**17.56**	**4.93**	**5.10**	**29.50**	**0.65**	**0.60**	**0.01**	**3.48**
NDXAM	**738**	**27.82**	**13.50**	**5.40**	**61.00**	**18.67**	**5.03**	**6.40**	**30.80**	**0.70**	**0.64**	**0.01**	**2.90**
Overall	**2411**	**26.39**	**13.57**	**5.10**	**72.20**	**18.16**	**4.50**	**4.50**	**33.98**	**0.66**	**0.64**	**0.01**	**3.99**

Note: N-sample size (number of trees), STD-standard deviation, Min.-minimum, Max.-maximum

Table 4. *Volume equations used for model fitting and model validation*

Model code	Model of Volume Equations	Number of parameters
1	$V = aD^b$	2
2	$V = aD^b H^c$	3
3	$V = a(D^2 H)^b$	2
4	$V = a + bD^2 H$	2
5	$V = aD^2 H$	1
6	$V = a1 + aD^b H^c$	3

V = over-bark stem volume (in m³), D = diameter at breast height (in cm), H = total tree height (in m), and a, b, c, and are parameters to be estimated from the data

Across all species and volume parameters, a simple power function proved to provide both consistent and accurate results (model code 2) and was thus selected as the basic model form for development (*Equation 1*):

$$V = aD^b H^c \tag{Eq.1}$$

Where V is volume in cubic meter (m^3), D is the tree diameter at breast height (DBH) (cm), H is the total height, a is the asymptotic parameter, b is the rate parameter, and c is the shape parameter. *Equation 1* was fit to (1) the Overall data (2411 Dahurian larch (*Larix gmelinii* Rupr) species of tree), (2) the region 1: The northwest of the northern slope of Yilehuli Mountains (NWYLHLM), (3) the region 2: The southeast of the northern slope of Yilehuli Mountains (SAYLHLM), (4) the region 3: The eastern slope of northern part of Daxing'an Mountains (NDXAM) and (5) each of the ten subregion separately. The PROC NLIN procedure in the Statistical Analysis System (SAS Institute, Inc. 1999) was utilized to estimate the model parameters and model statistics.

To test for differences between the overall model (region wide) and each ecoregion, a nonlinear extra sum of squares procedure was employed (Neter et al., 1996). This procedure involves the use of two indicator variables (k=2) who are needed in *Equation 2* for three regions in the full model form, while the reduced model form is represented by a three parameters model representing the volume relationship across all ecoregions (region wide). The full model form uses an indicator variable (k=2) approach to represent the tree regions.

- If region = NWYLHLM, $z_1 = 1$, $z_2 = 0$.
- If region = SAYLHLM, $z_1 = 0$, $z_2 = 1$.
- If region = NDXAM, $z_1 = 0$, $z_2 = 0$.

And to compare the difference of the volume relationships among different ecoregions, the non-linear extra sum of squares method was used (Bates and Watts, 1988; Huang et al., 2000). The method requires the fitting of full and reduced models.

The full model corresponds to different sets of parameters for each of the ecoregions involved, while the reduced model corresponds to the same set of parameters for all ecoregions.

Indicator (dummy) variables were used to facilitate the hypothesis tests as follows:

Test on the Overall model

In order to examine the sufficiency of the Overall model for the ten subregions, the full model for Darurian larch species and volume parameter was tested independently and is of the form:

$$V = \left(\alpha + \sum_{i=1}^{k} \alpha_i r_i\right) D^{\left(\beta + \sum_{i=1}^{k} \beta_i r_i\right)} H^{\left(\sigma + \sum_{i=1}^{k} \sigma_i r_i\right)}$$

(Eq.2)

Where:
V = volume parameter tested,
ri = indicator variable for ecoregion ri, i = 1,...,4,
D = tree dbh (in),
α,β, σ = parameters to be estimated from the data.
k= the number of indicator variables
The full model form has 10 parameters and an error sum of squares (SSEF) with N-10 degrees of freedom (dfF) where N is the total number of sample trees for each volume parameter and species tested. The form of the reduced model is that of *Equation 1* having three parameters and an error sum of squares (SSER) with N-3 degrees of freedom (dfR). The full model test has the following null and alternative hypotheses for each of the one species tests:
H0: $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \beta_1 = \beta_2 = \beta_3 = \beta_4 = \sigma_1 = \sigma_2 = \sigma_3 = \sigma_4 = 0$ and
Ha: at least one parameter is not equal to 0.

Rejecting the null hypothesis would indicate that the volume-diameter relationship is not the same for all ecoregions. Failure to reject the null hypothesis would indicate that the reduced model form *Equation 2* would be adequate for all ecoregions. These tests were conducted independently for each of the major species investigated.

In addition to the overall ecoregion tests, similar tests were conducted for each of the ten pairwise ecoregions comparisons for each of one species and three volume units examined. The same indicator variable approach was applied to the specific ecoregion comparison tests having a full model.

Test on the region against regions

To compare the volume relationship between the region and regions, the following full model was used

$$V = (\alpha + \alpha_1 r_1)D^{(\beta+\beta_1 r_1)}H^{(\sigma+\sigma_1 r_1)}$$ (Eq.3)

Where all variables and parameters were as previously defined. The full model form has three parameters and an error sum of squares (SSE_F) with N-4 degrees of freedom (df_F). The reduced model form is that of *Equation 1*. For each species and volume type tested, the full model test has the following null and alternative hypotheses:

H_0: $\alpha_1 = \beta_1 = \sigma_1 = 0$
and
H_a: at least one parameter is not equal to 0.

Rejection of the null hypothesis would indicate that the volume-diameter relationship is not the same between the three ecoregions being tested. Failure to reject the null hypothesis would indicate that a single model (*Equation 1*) could be used for both ecoregions for that species and volume type. These tests were conducted independently for each species and volume type combination.

To evaluate the differences between the three regions

A total of three regions pairs were formulated to test the pairwise differences between the three regions. The three testing pairs require three full models that take the form of *Equation 2*, and three reduced models that take the form of *Equation 1*. For example, to test the difference between region NWYLHLM vs SAYLHLM, one indicator variable (k=1) can be defined: if subregion = NWYLHLM, Z1 =1; and if subregion = SAYLHLM, Z1 = 0. Similarly, the full model (*Equation 2*) has 6 estimable parameters. To evaluate the overall differences among subregions nine indicator variables (k=9) are needed in *Equation 3* for ten subregions. They are defined as follows:

- If subregion = SL, $z1 = 1$, all other $z_i = 0$.
- If subregion = JGDQ, $z2 = 1$, all other $z_i = 0$.
- If subregion = XL, $z3 = 1$, all other $z_i = 0$.
- If subregion = TH, $z4 = 1$, all other $z_i = 0$.
- If subregion = HZ, $z5 = 1$, all other $z_i = 0$.
- If subregion = SBZ, $z6 = 1$, all other $z_i = 0$.
- If subregion = HJY, $z7 = 1$, all other $z_i = 0$.
- If subregion = XLJ, $z8 = 1$, all other $z_i = 0$.
- If subregion = TQ, $z9 = 1$, all other $z_i = 0$.
- If subregion = AME, all other $z_i = 0$.

Test on the differences between the ten subregions

To test the pairwise differences between the ten subregions, a total of 45 subregions pairs for Dahurian larch species were formulated. The 45 testing pairs require 45 full models that take the form of *Equation 3*, and 45 reduced models that take the form of *Equation 1*. For example, to test the difference between subregions SL vs. XL, one indicator variable (k=1) can be defined: if subregions = SL, Z1 =1; and if subregion = XL, Z1 = 0. Similarly, the full model (*Equation 3*) has 6 estimable parameters. The reduced model takes the form of *Equation 1* with 3 parameters. The null and alternative hypotheses are:

H_0 (1): $a_1 = b_1 = C_1 = 0$

H_a (2): at the least one parameter is not equal to 0.

Rejecting the null hypothesis H_a (2) would indicate that there is a difference between the two subregions tested (subregion 1 (SL) and 2 (JGDQ) in this example).

Test statistic

The significance of the full and reduced model comparisons were based on an F-test of the form:

$$F = \frac{\dfrac{SSE_R - SSE_F}{df_R - df_F}}{\dfrac{SSE_F}{df_F}}$$

(Eq.4)

Where SSE_F is the error sum of squares of a full model with the degrees of freedom df_F, and SSE_R is the error sum of squares of a reduced model with the degrees of freedom df_R.

This test statistic is F-distributed for a non-linear model if the data used represent a large sample generally, the F-test is significant if the p-value for the test is less than 0.05.

In order to understand the consequences of inappropriate application of a volume model in different ecoregions, each of the ten models (global model, regional model 1 (NWYLHLM 1), regional model 2 (SAYLHLM 2), regional model 3 (NDXAM 3) (and 10 suregions models) was used to predict total volume of trees for each ecoregion.

And to quantify the magnitude of the prediction error when a specific region model is used, the average volume prediction error $(\bar{\varepsilon})$, the standard error of the prediction error (S_e), and the prediction bias as a percentage of average "real" volume (% bias) were calculated and defined as:

$$\bar{\varepsilon} = \frac{\sum_{i=1}^{m} (\hat{V}_i - V_i)}{m}$$

(Eq.5)

$$S_e = \sqrt{\frac{\sum_{i=1}^{m} (\varepsilon - \bar{\varepsilon})^2}{m-1}}$$

(Eq.6)

$$Bias(\%) = \bar{\varepsilon} \Big/ \bar{V} \times 100$$

(Eq.7)

Where:

m= number of tree volumes for each species and volume type,

V_i= data reported volume for tree i,

$\widehat{V_i}$= predicted volume for tree i,

$\overline{V}$= mean data reported volume for each species and volume type. In this case $\overline{V}$ is the mean of observed tree volume. The t-test is used to test the null hypothesis and demonstrate that the mean prediction error is zero by

$$t = \frac{\bar{e}}{S_{\bar{e}}/\sqrt{m}}$$ (Eq.8)

Large values for these measures of prediction error would indicate situations where selection of an ecoregion specific model may be justified. Parameter estimates for *Equation 1* were obtained using SAS Proc NLIN (SAS 2002) for each species, ecoregion and volume type combination.

Results and Discussion

Estimation and validation of volume equations

The initial volume model development was completed with six candidate model forms. For all of the models, the significance of each parameter was tested and proved before and after weighting. The estimated parameters and their MSE-model mean squared error of the weighted models are presented in *Table 5* who is the parth of the data who help us to form the model.

Table 5. *Parameter estimations and model MSE of volume function for the regional and ecoregional models*

Ecoregions	N	a	b	c	MSE
SL	214	0.0000565	1.6187	1.2814	0.00997
JGDQ	524	0.00009	1.8015	0.9107	0.0114
XL	190	0.000049	1.8909	1.0325	0.0189
TH	203	0.000070	1.7627	1.0569	0.00494
HZ	215	0.000052	2.2630	0.5791	0.00914
SBZ	237	0.000076	1.6698	1.1221	0.00588
HJY	128	0.000059	1.8506	0.9993	0.00690
XLJ	224	0.000077	1.4784	1.3240	0.00135
TQ	191	0.000041	1.6976	1.3134	0.0165
AME	285	0.000131	1.5815	1.0603	0.0217
NWYLHLM	**915**	**0.000157**	**1.7457**	**0.8174**	**0.0156**
SAYLHLM	**758**	**0.000051**	**1.8080**	**1.1025**	**0.0112**
NDXAM	**738**	**0.000080**	**1.7574**	**1.0044**	**0.0112**
Overal	**2411**	**0.000103**	**1.7658**	**0.9220**	**0.0134**

Note: N-sample size (number of tree), a, b, c-three parameters of volume model, MSE-model mean squared error

The volume equations were classified as two types of functions: volume = f(DBH) with an independent variable and volume = f(DBH, height) with two independent variables (*Table 4*) and the parameter estimates and fit statistics were computed by species and all equations (*Table 5*). For the volume equations with DBH, the parameter estimates were valid ($P < 0.05$), but the P-value of parameter b was not valid in Equation (model 1) for Dahurian larch. For this reason, Equation (model 1) failed to fit the model. Equation (model 3) had lower MSE than Equation (model 1).

For the volume equations with DBH and height, parameter estimates of Equations (model 4) to (model 6) were firstly evaluated, and all parameters were valid ($P < 0.05$). Compared to other equations, however, Equations (model 3) and (model 5), especially for Dahurian larch, had fit statistics with high MSE. Meanwhile, Equation (model 6) estimated using nonlinear regression analysis predicted the volume with negative number when DBH was smaller than 10 cm, which indicated the instability of parameter estimates. For these reasons, Equations (model 5) and (model 6) failed to be the best fit equations in this study.

Equation (model 2) has valid coefficients and good fit statistics for Dahurian larch species and is considered as the best model to predict the total volume. This result corresponded to previous studies, which concluded that Equation (model 2), the so-called combined-variable function, is the best fit model (Burkhart, 1977; Amateis and Burkhart, 1987; Sherrill et al., 2011; Lee et al., 2017).

Fitting the total volume equations to regional and subregional data

The sample sizes, parameter estimates and mean square errors (MSE) for all models are provided in *Table 5*.

The four regional models (Overall, Region 1: NWYLHLM, Region 2: SAYLHLM and Region 3: NDXAM) have similar model MSEs. The parameters of the Overall model were similar to those of the regional model 1, due to the fact that regional1 data were 37.9% of the Overall data set with similar DBH and volume ranges for Dahurian larch species (*Table 3*). Among the ten ecoregional models, the lowest MSE value was found in the subregion 8 (XLJ) with a value of 0.00135. The parameter estimates also varied among the ten subregions, indicating that each ecoregion may have a different volume relationship from others.

Comparison of the total volume equations among subregions and between regions

Tables 1 and 6 show the testing results for the differences in Volume relationships between the Overall regions and among the ten subregions. *Equation 2* was used to test if the Overall model using all data combined was sufficient for the ten subregions.

Table 6. F-test for testing difference between regions

Regions	n	Full Model			Reduced Model			Extra sum of squares		
		SSE_F	PF	df_F	SSE_R	PR	df_R	$df_R - df_F$	PF	P_VALUE
Overal	2411	30.9453	9	2402	32.3012	3	2408	6	17.5411	0.0000
NWYLHLM - SAYLHLM	1673	22.7335	6	1667	23.3796	3	1670	3	15.7934	<. 0.0001
NWYLHLM - NDXAM	1653	22.4578	6	1647	23.1683	3	1650	3	17.3701	<. 0.0001
SAYLHLM - NDXAM	1496	16.6992	6	1490	17.4872	3	1493	3	23.4346	<. 0.0001

Note: N – sample size, SSE_F – error sum of squares of the full model, df_F – degrees of freedom of SSE_F, SSE_R – error sum of squares of reduced model, df_R – degrees of freedom of SSE_R, P is the number of parameters

The null hypothesis H_0 (1) was rejected (p-value < 0.0001), indicating that the Overall model was not sufficient to describe the volume-diameter relationships for the ten suregions. The difference between the different regions was also tested using *Equation 3* and the null hypothesis H_0 (2) was rejected (p-value < 0.0001). This means

that the difference in Volume relationships between the three geographic regions was statistically significant. A pairwise test revealed that there were differences in Volume relationships among ten subregions, i.e., the F-values ranged from 0.81 to 94.63 and all p-values were not less than 0.0001 (*Table 1*). Each ecoregion appeared to have a unique tree Volume relationship.

Prediction errors of applying subregional equation to each subregion

According to the tests above, there were significant differences among the ten subregions in the volume relationships. Inappropriately "applying" a volume model in these ecoregions may result in prediction biases. To understand the consequences, all fourteen models (Overall model, regional 1 model, regional 2 model, regional 3 model and ten subregional models) were used to predict total tree volume for each given ecoregion individually. The mean prediction error ($\bar{\varepsilon}$), the prediction bias (%), and the t-test for testing the null hypothesis that the mean prediction error equals zero are shown in *Tables 7 to 8*.

Table 7. *Prediction error of applying the overall model for each subregions*

Subregions	N	$\bar{V}$(m)	$\hat{V}$(m)	$\bar{\varepsilon}$	S_e	Bias (%)
SL	214	0.8126	0.8312	-0.0186	0.1039	-2.2859
JGDQ	524	0.6584	0.6852	-0.0267	0.1069	-4.0588
XL	190	0.8727	0.8241	0.0487	0.1572	5.5797
TH	203	0.4517	0.4489	0.0278	0.0721	0.6154
HZ	215	0.4977	0.4776	0.0201	0.1262	4.0371
SBZ	237	0.6442	0.6776	-0.0334	0.0833	-5.1780
HJY	128	0.6170	0.6295	-0.0124	0.0839	-2.0033
XLJ	224	0.2401	0.2513	-0.0112	0.0559	-4.6555
TQ	191	0.6234	0.5973	0.0262	0.1377	4.1955
AME	285	1.0806	1.0891	-0.0843	0.1526	-0.7798

Note: N – sample size, $\bar{V}$ – average of observed tree volume, $\hat{V}$ – average of predicted tree volume from the model, S_e – standard deviation of prediction error

If the overall model was used to predict the volume of trees in each region, excessive or underestimated predictions were made for different ecoregions (The p values of the t tests would be less than 0.05 for each of the ten subregions). On average, the overall model under-estimated (i.e., positive Bias %) tree volume from 0.6% to 5.6% for subregions XL, TH and HZ, and over-estimated (i.e., negative Bias %) tree volume about -2% to -5.2% in subregions HJY and SBZ. In the three forest regions, the Overall model under-estimated tree volume in both subregion TQ (4,2%) to subregion XL (5,6%) (*Table 7*).

Prediction errors of applying regional and overall volume equations to each subregion

Applying the regional model to each subregion would result in similar patterns and magnitudes of the prediction errors as the Overall model. This model under-estimated tree volume in the NWYLHLM - TQ (1.6%). And over-estimated tree volumes in the NWYLHLM - XLJ (-9.8%), NWYLHLM - SBZ (-8.5%), NWYLHLM - JGDQ (-7.1%), NWYLHLM - HJY (-5.5%), NWYLHLM - SL (-4.2%), NWYLHLM - TH (-3.4%) and NWYLHLM - NWYLHLM (-1.8%).

Table 8. *Prediction errors of applying a specific subregion model to each subregion*

Subregions	N	$\bar{V}$(m)	$\hat{V}$(m)	$\bar{\varepsilon}$	S_e	Bias (%)
SL 1						
SL - SL	214	0.8127	0.8185	-0.0588	0.0993	-0.7233
SL - JGDQ	524	0.6585	0.6612	-0.0277	0.1132	-0.4206
SL - XL	190	0.8728	0.7898	0.0829	0.1766	9.5095
SL - TH	203	0.4517	0.4277	0.0241	0.0813	5.3276
SL - HZ	215	0.4977	0.4516	0.0461	0.1474	9.2601
SL - SBZ	237	0.6442	0.6474	-0.0314	0.0772	-0.4877
SL - HJY	128	0.6171	0.6015	0.0156	0.0937	2.5228
SL - XLJ	224	0.2401	0.2480	- 0.0791	0.0456	-3.2931
SL - TQ	191	0.6234	0.5799	0.0435	0.1418	6.9854
SL - AME	285	1.0807	1.0908	-0.0101	0.1535	-0.9366
JGDQ 2						
JGDQ - SL	214	0.8127	0.7989	0.0138	0.1056	1.6923
JGDQ - JGDQ	524	0.6585	0.6577	0.0073	0.1065	0.1106
JGDQ - XL	190	0.8728	0.7922	0.0805	0.1652	9.2266
JGDQ - TH	203	0.4517	0.4297	0.0221	0.0774	4.8805
JGDQ - HZ	215	0.4977	0.4568	0.0408	0.1329	8.2106
JGDQ - SBZ	237	0.6442	0.6505	-0.0632	0.0799	-0.9811
JGDQ - HJY	128	0.6171	0.6036	0.0135	0.0866	2.1839
JGDQ - XLJ	224	0.2401	0.2380	0.0213	0.0532	0.8886
JGDQ - TQ	191	0.6234	0.5732	0.0501	0.1449	8.0476
JGDQ - AME	285	1.0807	1.0477	0.0329	0.1535	3.0486
XL 3						
XL - SL	214	0.8127	0.8871	-0.0745	0.1389	-9.1656
XL - JGDQ	524	0.6585	0.7193	-0.0609	0.1349	-9.2456
XL - XL	190	0.8728	0.8692	0.0356	0.1367	0.4075
XL - TH	203	0.4517	0.4611	-0.0937	0.0775	-2.0747
XL - HZ	215	0.4977	0.4869	0.0107	0.1085	2.1590
XL - SBZ	237	0.6442	0.7085	-0.0643	0.1179	-9.9776
XL - HJY	128	0.6171	0.6540	-0.0369	0.0937	-5.9855
XL - XLJ	224	0.2401	0.2479	-0.0777	0.0836	-3.2342
XL - TQ	191	0.6234	0.6285	-0.0509	0.1328	-0.8180
XL - AME	285	1.0807	1.1776	-0.0969	0.2066	-8.9667
TH 4						
TH - SL	214	0.8127	0.8546	-0.0419	0.1088	-5.1609
TH - JGDQ	524	0.6585	0.6968	-0.0384	0.1128	-5.8291
TH - XL	190	0.8728	0.8378	0.0350	0.1483	4.0124
TH - TH	203	0.4517	0.4518	-0.0009	0.0699	-0.0221
TH - HZ	215	0.4977	0.4786	0.0191	0.1228	3.8285
TH - SBZ	237	0.6442	0.6864	-0.0421	0.0889	-6.5414
TH - HJY	128	0.6171	0.6364	-0.0193	0.0834	-3.1301
TH - XLJ	224	0.2401	0.2523	-0.0121	0.0625	-5.0579
TH - TQ	191	0.6234	0.6089	0.0145	0.1310	2.3321
TH - AME	285	1.0807	1.1296	-0.0489	0.1651	-4.5282
HZ 5						
HZ - SL	214	0.8127	0.8855	-0.0729	0.1795	-8.9669
HZ - JGDQ	524	0.6585	0.7303	-0.0718	0.1778	-10.908
HZ - XL	190	0.8728	0.8963	-0.0235	0.1678	-2.6975
HZ - TH	203	0.4517	0.4666	-0.0149	0.1154	-3.3011
HZ - HZ	215	0.4977	0.4948	0.0291	0.0952	0.5849
HZ - SBZ	237	0.6442	0.7277	-0.0834	0.1827	-12.949
HZ - HJY	128	0.6171	0.6667	-0.0496	0.1354	-8.0448
HZ - XLJ	224	0.2401	0.2279	0.0122	0.1201	5.0896
HZ - TQ	191	0.6234	0.6339	-0.0105	0.1628	-1.6775
HZ - AME	285	1.0807	1.1606	-0.0799	0.2524	-7.3936
SBZ 6						
SBZ - SL	214	0.8127	0.8107	.0.0200	0.1007	0.2466
SBZ - JGDQ	524	0.6585	0.6613	-0.0281	0.1092	-0.4263
SBZ - XL	190	0.8728	0.7918	0.0809	0.1745	9.2726
SBZ - TH	203	0.4517	0.4308	0.0209	0.0797	4.6327
SBZ - HZ	215	0.4977	0.4566	0.0411	0.1427	8.2544
SBZ - SBZ	237	0.6442	0.6504	-0.0614	0.0763	-0.9533

Subregions	N	$\bar{V}$(m)	$\hat{V}$(m)	$\bar{\varepsilon}$	S_e	Bias (%)
SBZ - HJY	128	0.6171	0.6046	0.0125	0.0909	2.0298
SBZ - XLJ	224	0.2401	0.2471	-0.0693	0.0463	-2.8855
SBZ - TQ	191	0.6234	0.5784	0.0450	0.1439	7.2243
SBZ - AME	285	1.0807	1.07152	0.0917	0.1501	0.8484
HJY 7						
HJY - SL	214	0.8127	0.82939	-0.0167	0.1082	-2.0599
HJY - JGDQ	524	0.6585	0.67592	-0.0175	0.1111	-2.6531
HJY - XL	190	0.8728	0.81559	0.0572	0.1478	6.5513
HJY - TH	203	0.4517	0.43620	0.0155	0.0713	3.4374
HJY - HZ	215	0.4977	0.46176	0.0359	0.1223	7.2192
HJY - SBZ	237	0.6442	0.66660	-0.0224	0.0891	-3.4731
HJY - HJY	128	0.6171	0.61655	0.0053	0.0824	0.0862
HJY - XLJ	224	0.2401	0.23754	0.0259	0.0646	1.0803
HJY - TQ	191	0.6234	0.59013	0.0333	0.1338	5.3383
HJY - AME	285	1.0807	1.09658	-0.0159	0.1671	-1.4709
XLJ 8						
XLJ - SL	214	0.8127	0.77025	0.0424	0.1178	5.2179
XLJ - JGDQ	524	0.6585	0.62573	0.0327	0.1309	4.9692
XLJ - XL	190	0.8728	0.74293	0.1298	0.2206	14.876
XLJ - TH	203	0.4517	0.40948	0.0423	0.1082	9.3531
XLJ - HZ	215	0.4977	0.43355	0.0641	0.1745	12.888
XLJ - SBZ	237	0.6442	0.61228	0.0319	0.0938	4.9585
XLJ - HJY	128	0.6171	0.57157	0.0455	0.1215	7.3756
XLJ - XLJ	224	0.2401	0.24797	-0.0784	0.0357	-3.2643
XLJ - TQ	191	0.6234	0.54921	0.0742	0.1725	11.9033
XLJ - AME	285	1.0807	1.02260	0.0581	0.1587	5.3743
TQ 9						
TQ - SL	214	0.8127	0.8856	-0.0729	0.1268	-8.9781
TQ - JGDQ	524	0.6585	0.7101	-0.0516	0.1280	-7.8375
TQ - XL	190	0.8728	0.8506	0.0221	0.1462	2.5369
TQ - TH	203	0.4517	0.4543	-.0.0261	0.0729	-0.5779
TQ - HZ	215	0.4977	0.4781	0.0196	0.1258	3.9439
TQ - SBZ	237	0.6442	0.6941	-0.0499	0.0954	-7.7461
TQ - HJY	128	0.6171	0.6428	-0.0257	0.0891	-4.1653
TQ - XLJ	224	0.2401	0.2569	-0.0167	0.0675	-6.9841
TQ - TQ	191	0.6234	0.6231	0.0033	0.1279	0.0525
TQ - AME	285	1.0807	1.1869	-0.1062	0.1951	-9.8283
AME 10						
AME - SL	214	0.8127	0.8327	-0.0199	0.1023	-2.4606
AME - JGDQ	524	0.6585	0.6865	-0.0280	0.1096	-4.2542
AME - XL	190	0.8728	0.8189	0.0538	0.1801	6.1619
AME - TH	203	0.4517	0.4539	-0.0216	0.0794	-0.4791
AME - HZ	215	0.4977	0.4831	0.0146	0.1425	2.9319
AME - SBZ	237	0.6442	0.6766	-0.0324	0.0767	-5.0306
AME - HJY	128	0.6171	0.6317	-0.0146	0.0928	-2.3688
AME - XLJ	224	0.2401	0.2681	-0.0279	0.0432	-11.634
AME - TQ	191	0.6234	0.5999	0.0236	0.1449	3.7786
AME - AME	285	1.0807	1.0914	-0.0107	0.1464	-0.9879

Note: N – sample size, $\bar{V}$ – average of observed tree volume, $\hat{V}$ – average of predicted tree volume from the model, S_e – standard deviation of prediction error

Again, larger under-estimations occurred in the NWYLHLM - XL (2.9%) (*Table 9*). This was expected because the estimated parameters of the Overall model were close to those of the three regional models (*Table 5*).

On the other hand, the regional model has not been able to overestimate the volume of trees in the ten ecoregions of the different regions (*Table 9*).

When the ten ecoregional models were applied to each ecoregion, they generally performed well in the ecoregions in which the models were developed (p-values for the t-

tests > 0.05 for testing that the mean prediction error equals zero) and the models produced significant prediction errors. The prediction biases ranged from -3.3% to 9.5% in Model 1 (SL1), -0.9% to 9.2% in Model 2 (JGDQ2), -9.9% to 2.1% in Model 3 (XL3), -6.5% to 4% in Model 4 (TH4), -12.9% to 5.1% in Model 5 (HZ5), -2.8% to 9.3% in Model 6 (SBZ6), -3.5% to 7.2% in Model 7 (HJY7), -3.3% to 14.9% in Model 8 (XLJ8), -9.8% to 3.9% in Model 9 (TQ9) and -11.6 to 6.2% in Model 10 (AME10) (*Table 8*).

Table 9. Prediction errors of applying regional models to each subregion

Reg – Sub	N	$\bar{V}$(m)	$\hat{V}$(m)	$\bar{\varepsilon}$	S_e	Bias (%)
NWYLHLM - SL	214	0.8126	0.8467	-0.0341	0.1062	-4.1954
NWYLHLM - JGDQ	524	0.6584	0.7048	-0.0464	0.1076	-7.0473
NWYLHLM - XL	190	0.8727	0.8470	0.0257	0.1569	2.9456
NWYLHLM - TH	203	0.4517	0.4668	-0.0151	0.0722	-3.3517
NWYLHLM - HZ	215	0.4977	0.4984	-0.0077	0.1227	-0.1553
NWYLHLM - SBZ	237	0.6442	0.6992	-0.0549	0.0871	-8.5312
NWYLHLM - HJY	128	0.6171	0.6510	-0.0339	0.0850	-5.5072
NWYLHLM - XLJ	224	0.2401	0.2636	-0.0235	0.0565	-9.7865
NWYLHLM - TQ	191	0.6234	0.6132	0.0101	0.1384	1.6296
NWYLHLM - AME	285	1.0807	1.1005	-0.0199	0.1503	-1.8412
SAYLHLM - SL	214	0.8126	0.8491	-0.0365	0.1128	-4.4898
SAYLHLM - JGDQ	524	0.6584	0.6879	-0.0295	0.1158	-4.4792
SAYLHLM - XL	190	0.8727	0.8284	0.0443	0.1453	5.0817
SAYLHLM - TH	203	0.4517	0.4423	0.0936	0.0704	2.0724
SAYLHLM - HZ	215	0.4977	0.4672	0.0304	0.1220	6.1162
SAYLHLM - SBZ	237	0.6442	0.6765	-0.0322	0.0913	-5.0072
SAYLHLM - HJY	128	0.6171	0.6257	-0.0868	0.0833	-1.4062
SAYLHLM - XLJ	224	0.2401	0.2433	-0.0320	0.0661	-1.3335
SAYLHLM - TQ	191	0.6234	0.6016	0.0217	0.1302	3.4885
SAYLHLM - AME	285	1.0807	1.1281	-0.0474	0.1757	-4.3865
NDXAM - SL	214	0.8126	0.8117	0.0094	0.1023	0.1158
NDXAM - JGDQ	524	0.6584	0.6649	-0.0647	0.1067	-0.9832
NDXAM - XL	190	0.8727	0.7993	0.0734	0.1639	8.4153
NDXAM - TH	203	0.4517	0.4332	0.0185	0.0755	4.1036
NDXAM - HZ	215	0.4977	0.4597	0.0379	0.1337	7.6241
NDXAM - SBZ	237	0.6442	0.6559	-0.0117	0.0787	-1.8260
NDXAM - HJY	128	0.6171	0.6088	0.0826	0.0858	1.3380
NDXAM - XLJ	224	0.2401	0.2426	-0.0247	0.0524	-1.0274
NDXAM - TQ	191	0.6234	0.5804	0.0429	0.1411	6.8916
NDXAM - AME	285	1.0807	1.0689	0.0117	0.1526	1.0849

Note: N – sample size, $\bar{V}$ – average of observed tree volume, $\hat{V}$ – average of predicted tree volume from the model, S_e – standard deviation of prediction error

These results are expected because different ecoregions have large differences in climatic, soil and ecological conditions. As Peng et al. (2004) and Rivas et al. (2007) indicated, part of the differences in dbh and volume models among tree species may be caused by differences in genetics, growth conditions and site quality. Recent studies with different species also indicated differences in ecoregion-based height diameter models (Huang et al., 1999; Peng et al., 2004; Brooks and Wiant, 2005). For our case, the updated Ecoregion classification of Dahurian larch region (*Figure 1*) provides comprehensive information on changes in ecoregion-species relationships along the ecological and macroclimatic gradients of different regions. These ecoregions are characterized by large climatic patterns (temperature and precipitation, for example), soil moisture and nutrient

regimes, and succession and vegetation types (Hills, 1959, 1960; CEC, 1997; ELC Working Group, 2000).

In general, Dahurian larch species are larger in forest subregions (JGDQ and AME), mainly due to climate, abundant rainfall and the longer growing season. For example, the average annual temperatures of the different ecoregions are between -1°C to -2.8°C and -28°C to 20°C, respectively in Daxing'an Mountains region. Moreover the average air temperatures in January and July are about -28°C and 20°C, respectively while the mean annual precipitation is relatively higher in Daxing'an Mountains region (500 to 750 mm). More detailed analysis of climatic characteristics for each ecoregion can be found in Dong et al. (2014). In addition, local ecological conditions (vegetation type) also affect species growth and productivity in these ecoregions. For example, in the ten subregions studied, it has been shown that the stand of the Dahurian larch species is the most dominant in relation to others species. The latter which is our study species comprises a total of 2411 individuals for all ten subregions compared to a total of 1324 trees in all ten subregions for *Betula platyphylla* species (BH), 445 trees for Siberian larch (SY) species, 1294 trees for Korean pine (ZZS). Subregion 2 (JGDQ) has the largest number of trees (524 trees for Dahurian larch species only).

Therefore, the combination of all climatic, environmental and vegetation factors plays an important role in determining the different volume relationships among ten subregions.

These results are consistent with the recent findings reported by Huang (1999) quoted by Liunjun et al. (2002) and Huang et al. (2000) for Lodgepole pine and white spruce in the boreal forests of Alberta.

Illustration of different graphics in Figure 2

The different figures below illustrate average prediction errors on diameter classes of 5 cm for the ten subregions. Clearly, the overall model's negative biases in some subregions were due to overestimation for large trees. The larger positive model biases in the two subregions (XL (second line on the left) and TQ (fifth line on the left)) are due to the underestimation of all trees by the global model.

This limitation was initial pointed by Thalmann (1965), where advanced evaluation methods based on confirmed technical bases are necessary. These graphs show the mean prediction errors (m^3) in diameter classes of 5 cm when the global model is applied to predict the volume of trees in each of the ten subregions.

Solid lines with solid symbols: diamond, square, triangle, asterisk and circle are used to represent subregions:

- First line (SL and JGDQ: 3E (circles), 3W (squares), 2S (triangles), 3S (asterisks) and 3W (circles)).
- Second line (XL and TH: 3E (circles), 3W (asterisk), 3W (squares), 3S (triangles).
- Third line (HZ and SBZ: 3E (triangles), 3W (triangles), 3W (circles), 3W (squares), 4W (asterisks), 2W (circles).
- Fourth line (HJY and XLJ: 3E (circles), 2W (circles), 4W (asterisks), 1W (squares), 1S (rod).
- Fifth line (TQ and AME: 2E (rods), 2E (circles), 2W (diamond), 5w (asterisks), 3W (squares).

Therefore, trends in prediction errors in the 5 cm of diameter classes were very similar for some regional model and the overall model.
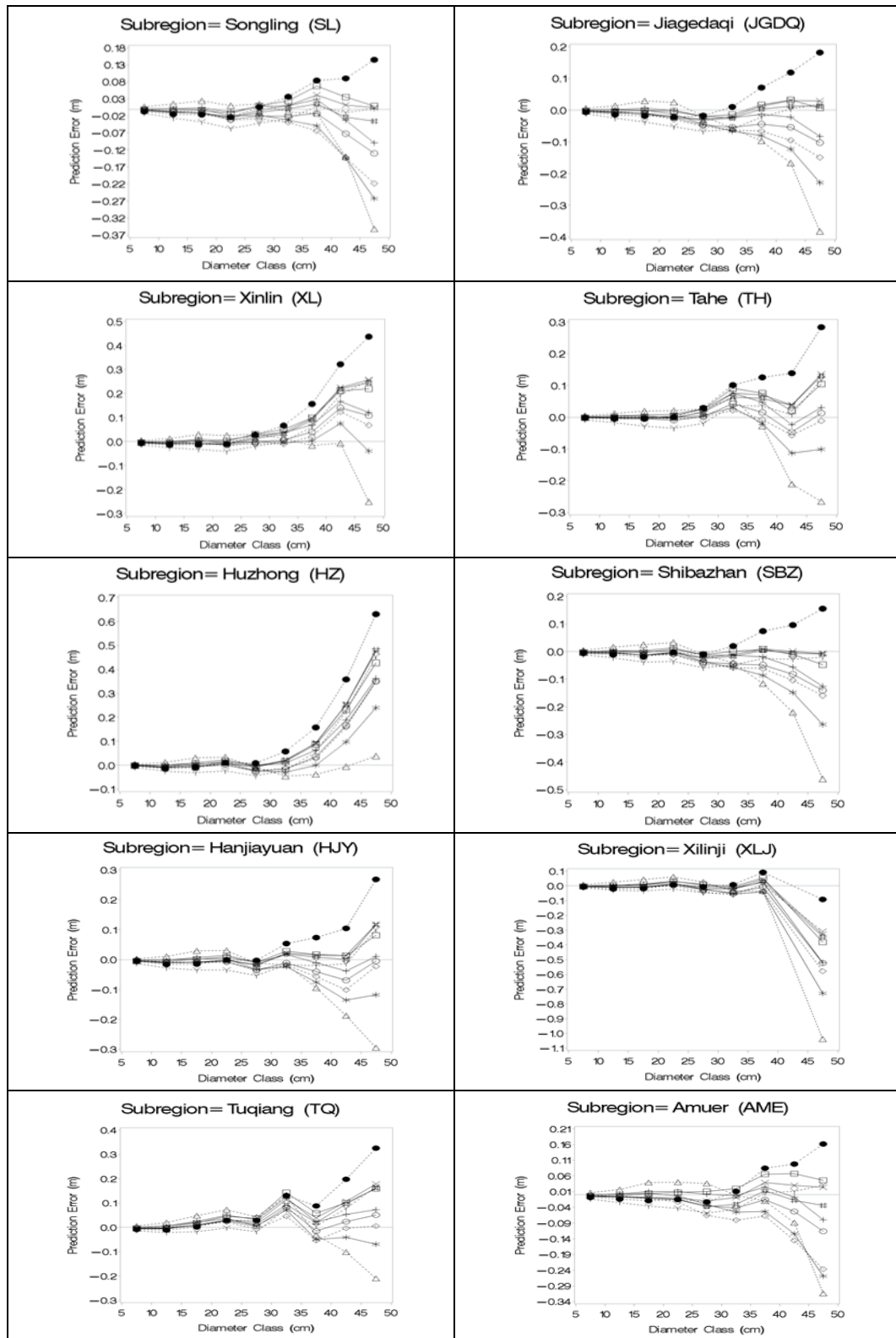
Figure 2. *Average prediction errors (m) across 5-cm diameter classes when each subregional model is applied to predict total volume in each of the ten subregions with symbols: SL (hash), JGDQ (square), XL (star),TH (circle), SBZ (X), HJY (plus), HZ (triangle), XLJ (dot), TQ (diamond) and AME (Y)*

On the other hand, other regional models overestimated tree volume (e.g. % bias) in some subregions of different regions including: HJY and SBZ and in others, underestimated, for example: XL, TH, HZ and finally TQ and XL. Some figures indicate that negative model biases result from overestimation of all trees in diameter classes of different subregions. The relatively high value for some subregions can be explained by the morphology of the Dahurian larch stem in the temperate mixed forests of the study regions.

Conclusion

With total tree volume equations being developed in the study for economically important hardwood species from northeastern China (Daxing'an Mountain region /Heilongjiang province), these equations were based on a simple function and adjusted to the forest inventory data volumes of three mountainous regions, including the north-west of the North Slope of the Yilehuli Mountains, the southeast of the North Slope of the Yilehuli Mountains, and the northeastern slope of the Daxin'an Mountains. Using dummy variables and an additional sum of squares procedure, the volume - diameter relationship was not the same for all ecoregions when tested by species and volume type. For ecoregion-specific comparisons, statistical differences in total volume were found across many ecoregions, which can be explained in part by the longitudinal, latitudinal, and elevation properties of differences between soils, the length of the season of growth and annual rainfall rate in the regions, which leads us to prove the significance of the relationship between the volume and diameter of the trees of the Dahurian larch tree species studied.

The main underlying factor is probably the differences in height-diameter relationships reported for the same ecoregions in northeastern China. Mean bias varied from -3.3% to 9.5% for (SL1), -0.9% to 9.2% for (JGDQ2), -9.9% to 2.1% for (XL3), -6.5% to 4% for (TH4), -12.9% to 5.1% for (HZ5), -2.8% to 9.3% for (SBZ6), -3.5% to 7.2% for (HJY7), -3.3% to 14.9% for (XLJ8), -9.8% to 3.9% for (TQ9) and -11.6% to 6.2% for (AME10) for the marketable value of timber in the three regions. Local tree volume equations specific to species and ecoregions are presented and can be easily used in wood intrusion situations. However, there are clear variations in volume relationships for Dahurian larch species in the ten subregions of the three regions studied in northeastern China (Heilongjiang Province). Inappropriate application of provincial or regional volume models to these different subregions can lead to significant errors in estimating tree height and total volumes. Ecoregion-based volume models can actually provide more accurate predictions of tree volume and height than models developed at the provincial or regional level.

REFERENCES

[1]   Akinsanmi, F. A., Akindélé, S. O. (1995): Teak productivity in relation to soil conditions: A re-assessment of teak plantations in the dry high forest area of south-western Nigeria. – Nigerian Journal of Forestry 24: 7-10.

[2]   Amateis, R. L., Burkhart, H. E. (1987): Tree volume and taper of loblolly pine varies by stand origin. – South J Appl For. 11: 185-189.

[3]   Barrio, M., Diéguez-Aranda, U., Castedo-Dorado, F., Álvarez, J. G., Kv, G. (2007): Merchantable volume system for pediculate oak in northwestern Spain. – Ann for Sci 64: 511-520. [CrossRefGoogleScholar].

[4]   Bates, D. M., Watts, D. G. (1988): Nonlinear regression analysis and its applications. – Wiley, New York.

[5]   Bontemps, J. D., Longuetaud, F., Franceschini, T., Charru, M., Constant, T. (2012): Estimation of forest biomass and productivity in the face of environmental change. – AgroParisTech, ENGREF, UMR 1092 INRA / AgroParisTech Forest-Wood Resources Study Laboratory (LERFoB), Street 14 Girardet, F-54000 Nancy 14p.

[6]   Brooks, J. R., Wiant, H. V. Jr. (2004): A simple technique for estimating cubic volume yields. – Forest Ecology and Management 203: 373-380.

[7]   Brooks, J. R., Wiant, H. V. Jr. (2005): Evaluating ecoregion based height_diameter relationship of five economically important Appalachian hardwood species in West Virginia. – In: McRoberts, R. E., Reams, G. A., Van Deusen, P. C., McWilliams, W. H. (eds.) Proceedings of the seventh annual forest inventory and analysis symposium; October 3-6, 2005; Portland, ME. Gen. Tech. Rep. WO-77. Washington, DC: U.S. Department of Agriculture, Forest Service: 237-242.

[8]   Brooks, J. R., Jiang, L., Ozçelik, R. (2008): Compatible stem volume and taper equations for Brutian pine, Cedar of Lebanon, and Cilicica fir in Turkey. – Forest Ecology and Management 256(1): 147-151.

[9]   Burkhart, H. E. (1977): Cubic-foot volume of loblolly pine to any merchantable top limit. – South J Appl For. 1: 7-9.

[10]   CEC (Commission for Environmental Cooperation). (1997): Ecological Regions of North America: Toward a common perspective. – CEC, Montreal. 71 p and maps.

[11]   Corral-Rivas, J. J., Barrio-Anta, M., Aguirre-Calderon, O. A., Dieguez-Aranda, U. (2007): Use of stump diameter to estimate diameter at breast height and tree volume for major pine species in El Salto, Durango (Mexico). – Forestry 80: 29-40.

[12]   CRPF FOGEFOR d, Limousin. (2011): The cubing of standing trees. 17 p.

[13]   Dong, L.-H., Zhang, L., Li, F. (2014): Developing additive system of biomas equations for nine hardwood species in Norteast China. – Forest Ecology and Management 329: 306-317.

[14]   Dong, L., Liu, Z., Bettinger, P. (2016): Nonlinear mixed-effects branch diameter and length models for natural Dahurian larch (Larix gmelini) forest in northeast China. – Trees 30(4): 1191-1206.

[15]   ELC Working Group. (2000): The Ecoregions and Ecodistricts of Ontario. – Ecological Land Classification Program, Ont. Min. Nat. Resour., Sault Ste Marie, ON.

[16]   Fonton, N. H., Glele Kakai, R., Rondeux, J. (2002): Étude dendrométrique d'Acacia auriculiformis A. Cunn. ex Benth. en mélange sur vertisol au Bénin. [Dendrometric study of Acacia auriculiformis A. Cunn. Ex Benth. in mixture on vertisol in Benin]. – Biotechnologie, Agronomie, Société et Environnement 6(1): 29-37.

[17]   Fonton, N. H., Yabi, C. C., Dovonon, J. Z. D., Adoko, F. K., Titilayo. D. (2009): Tree Shaft Volume Modeling for Sustainable Management of Sudanese Forest Ecosystems. 100 p.

[18]   Fonton, H. N., Dah-Dovonon, J. Z., Adoko, F. K., Dotchamou, T. (2009): Modélisation du volume du fût d'arbre pour une gestion durable des écosystèmes forestiers soudaniens.

[Modelling of tree stem volume for sustainable management of sudanian forest ecosystems]. – Bois et Forêt des Tropiques 300(2): 95-100.

[19]   Hills, G. A. (1959): A Ready Reference to the Description of the Land of Ontario and its Productivity. – Preliminary Report. Ontario Department of Lands and Forests, Division of Research, Maple. 142p.

[20]   Hills, G. A. (1960): Comparison of forest ecosystems (vegetation and soil) in different climatic zones. – Silva Fennica 105: 33-39.

[21]   Honer, T. G. (1967): Standard volume tables and merchantable conversion factors for the commercial tree species of central and eastern Canada. – Can. Dept. Forestry Rural Devel., For. Mgmt. Res. And Serv. Inst. Info. Rep. FMR-X-5.

[22]   Huang, S., Titus, S. J., Wiens, D. P. (1999): Comparison of nonlinear height_diameter functions for major Alberta tree species. – Canadian Journal of Forest Research 22: 1297-1304.

[23]   Huang, S., Price, D., Morgan, D., Peck, K. (2000): Kozak's variable-exponent taper equation regionalized for white spruce in Alberta. – West J Appl For 15(2): 75-85. [CrossRefGoogle Scholar].

[24]   Huang, S., Price, D., Titus, S. J. (2000): Development of ecologicalbased height–diameter models for white spruce in boreal forests. – For. Ecol. Manage. 129: 125-141. [CrossRefGoogleScholar].

[25]   Jia, B., Zhou, G. (2018): Growth characteristics of natural and planted Dahurian larch in northeast China. – Earth System Science Data Discussions 6p.

[26]   Jiang, L., Brooks, J. R., Wang, J. (2004): Compatible taper and volume equations for yellow-poplar in West Virginia. – For. Ecol. And Mgmt. 213: 399-409. [CrossRefGoogle Scholar].

[27]   Jin, Q., Valsta, L., Heliövaara, K., Li, J., Luo, Y., Shi, J. (2015): Effects of Catastrophic Insect Outbreaks on the Harvesting Solutions of Dahurian Larch Plantations. – International Journal of Forestry Research, 12p.

[28]   Laumans, P. (1991): Tables de cubage à deux entrées pour le volume bois sciable du Teck (Tectona grandis L. f.) au Sud Bénin. – Wittelsbacherstr 11, D-8016 Feldkirchen, 18p.

[29]   Lee, D., Seo, Y., Choi, J. (2017): Estimation and validation of stem volume equations for Pinus *densiflora*, *Pinus koraiensis*, and *Larix kaempferi* in South Korea. – Forest Science and Technology 13(2): 77-82.

[30]   Neter, J., Kutner, M. H., Nachtsheim, C. J., Wasserman, W. (1996): Applied linear statistical models. – McGraw-Hill, New York. 1048p.

[31]   Newnham, R. (1992): Variable-form taper functions for four Alberta tree species. – Can J For Res 22: 210-223. [CrossRefGoogleScholar].

[32]   Özcelik, R., Karatepe, Y., Gürlevik, N., Canellas, I., Crecente-Campo, F. (2016): Development of ecoregion-based merchantable volume systems for *Pinus brutia* Ten and *Pinus nigra* Arnold in southern Turkey. – Journal of Forestry Research 27(1): 101-117.

[33]   Palm, R. (1982): Influence of the volume formula and the length of the ridges on the determination of the volume of felled trees. – Annals of Forest Science 39: 231-238.

[34]   Palm, R., Iemma, A. F. (2002): Application conditions and variable transformations in linear regression. – Gembloux Belgium, University Faculty of Agricultural Sciences, Statistics and Computer Science Notes, 34p.

[35]   Pauwels, D., Rondeux, J. (1999): Tarifs de cubage pour les petits bois de melèze (Larix sp.) en Ardenne. – Les Cahiers Forestiers de Gembloux, 11p.

[36]   Peng, C., Zhang, L., Zhou, X., Dang, Q., Huang, S. (2004): Developing and evaluating tree height_diameter models at three geographic scales for black spruce in Ontario. – Northern Journal of Applied Forestry 21: 83-92.

[37]   Rondeux, J. (1999): The measurement of trees and forest stands. – Gembloux, Belgium, The Agronomic Press of Gembloux, 2nd edition, 521p.

[38]   SAS Institute Inc. (2002): SAS/STAT Online User's Guide, Version 9.0. – SAS Institute Inc., Cary, NC.

[39] Shailer, S., Kershaw, J. A., Zundel, P. (1998): Comparison of total volume equations for use in southwestern New Brunswick. – Research Report perpared for Georgia Pacific (The Timber Company), St. Croix district.

[40] Sherrill, J. R., Bullock, B. P., Mullin, T. J., McKeand, S. E., Purnell, R. C. (2011): Total and merchantable stem volume equations for midrotation loblolly pine (Pinus taeda L.). – South J Appl For. 35: 105-108.

[41] Thibaut, A., Rondeux, J., Claessens, H. (1998): Tarifs de cubage pour l'Aulne glutineux [Alnus glutineux (L.) Gaertn.] en Belgique méridionale. – Revue Forestière Française, L-3: 244-250.

[42] Tsega, M., Guadie, A., Teffera, Z. L., Belayneh, Y., Niu, D. (2018): Development and validation of a stem volume equation for Cupressus lusitanica in Gergeda Forest, Ethiopia. – Southern Forests: a Journal of Forest Science 81(1).

[43] Zhang, L., Peng, C., Huang, S., Zhou, X. (2002): Development and evaluation of ecoregion-based jack pine height-diameter models for Ontario. – Forestry Chronicle 78(4): 530-538.