Estimation and validation of stem volume equations for *Pinus densiflora*, *Pinus koraiensis*, and *Larix kaempferi* in South Korea

Daesung Lee*, Yeongwan Seo† and Jungkee Choi*

*Department of Forest Management, College of Forest and Environmental Sciences, Kangwon National University, Chuncheon 24345 Republic of Korea; †Institute of Forest Science, College of Forest and Environmental Sciences, Kangwon National University, Chuncheon 24345 Republic of Korea*

**ABSTRACT**

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. These models revealed higher accuracy when compared with previous studies. Both equations are considered to be easily used in the field.

**Introduction**

In Korea, Korean red pine (*Pinus densiflora* Siebold & Zucc.), Korean white pine (*Pinus koraiensis* Siebold & Zucc.), and Japanese larch (*Larix kaempferi* (Lamb.) Carrière) have been widely used as lumber, pulp, and other wood materials (Korea Forest Research Institute 2012a, 2012b, 2012c). These trees are important commercial species and have been intensively planted in South Korea since the 1960s (Korea Forest Service 2016). At present, these species cover large forest areas of 1,562,843 ha, 170,905 ha, and 272,800 ha, respectively, with annually planting of more than 7000 ha in total (Korea Forest Service 2016). Especially, the forest area larger than 52% of 2,006,548 ha, total area of the three species in South Korea, are covered by these species with 1,055,981 ha in Gangwon, Gyeonggi, and North Gyeongsang provinces, and so more comprehensive and intensive forest resource management is needed for *P. densiflora*, *P. koraiensis*, and *L. kaempferi* (Korea Forest Service 2016).

For forest resource management, forest stock is basically represented by stem volume, and volume equations are widely applied using diameter at breast height (DBH) and height to calculate stem volume (Avery and Burkhart 2002). Various volume equations have been studied by forest biometricians all over the world (Schumacher and Hall 1933; Spurr 1952; Honer 1965; Baskerville 1972). In particular, the stem volume equations of a species have been subdivided according to regional characteristics for accurate commercial use in other developed countries (Beck 1963; Myers 1972; Queen and Pienaar 1977; Amateis and Burkhart 1987; Clark and Saucier 1990).

In Korea, stem volume models have been developed for *P. densiflora*, *P. koraiensis*, and *L. kaempferi* by some researchers in the past. The Korea Forest Research Institute (2012) has developed stem volume equations using data from the whole country. Stem volume equations for *P. densiflora* of Gangwon province and Uljin region were studied by Kim et al. (1994) and Kim et al. (2001), respectively. Also, the models for *P. koraiensis* of Hongcheon and Gapyeong regions were studied by Shin (1994). The volume equations for *L. kaempferi* were studied by Kim and Choung (1988) with data from Gangwon and Gyeonggi provinces, by Shin et al. (1996) with data from Hongcheon region, and by Jeon et al. (2007) with data from Jinan region. However, the models were mostly developed with samples from young forests, and were limited to some local use. Therefore, new volume models are needed for the precise and suitable volume prediction on mature large trees of these commercially important species.

The objectives of this study were to provide new parameter estimates of total stem volumes, to compare the models with previous studies for validation, and to verify the best stem volume equation for *P. densiflora*, *P. koraiensis*, and *L. kaempferi* in Gangwon, Gyeonggi, and North Gyeongsang provinces of South Korea.

**Materials and methods**

**Data**

The data available for this study were collected from *L. kaempferi*, *P. koraiensis*, and *P. densiflora* stands located in Gangwon, North Gyeongsang, and Gyeonggi provinces of South Korea (Figure 1). In each site, one standard tree was collected as a sample tree and the total number of sample trees was 130: 45 trees for *L. kaempferi*, 47 trees for *P. koraiensis*, and 38 trees for *P. densiflora*. The mean of sample trees by species was 31.2, 29.8, and 27.8 cm in DBH and 17.4, 17.2, and 22.8 m in height, respectively. Characteristics of the 130 sample trees are summarized in Table 1.

The standard trees were cut down at 0.2 m above the ground, and first and second wood discs were collected from 0.2 m and 1.2 m, respectively. After 1.2 m, wood discs were collected at regular 2 m intervals: the specific procedure and method were reported by Lee and Choi (2016) in detail.
After collecting the wood discs, total stem volume of each tree was computed by using Smalian’s formula and by adding the top volume after last wood disc, assuming that the shape of top section was a cone (Knoebel et al. 1984; Avery and Burkhart 2002).

**Volume equations and evaluation**

Various volume equation forms were tried for the best model fit. The volume equations were comprised of DBH and total height as independent variables because of high suitability. These equations have been previously applied for tree volume models in the field of forestry (Schumacher and Hall 1933; Spurr 1952; Beck 1963; Honer 1965; Baskerville 1972; Burkhart 1977). Selected eight volume equations and characteristics are summarized in Table 2. The volume equations were classified as two types by the independent variables: volume = f(DBH) and volume = f(DBH, height).

Parameters were estimated by regression analysis using the PROC REG and PROC NLIN procedures on SAS 9.4 software (SAS Institute Inc. 2013). For model validation, coefficient of determination (R²), root mean square error (RMSE), mean deviation (MD), and mean absolute deviation (MAD) were used to evaluate and determine the best model (Lee et al. 2014; Park et al. 2015). The equations are summarized as follows:

\[
R^2 = 1 - \left[ \frac{\sum_{i=1}^{n} (V_i - \bar{V}_i)^2}{\sum_{i=1}^{n} (V_i - \bar{V})^2} \right]
\]

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{n} (V_i - \bar{V}_i)^2}{n}}
\]

\[
MD = \frac{\sum_{i=1}^{n} (V_i - \bar{V}_i)}{n}
\]

\[
MAD = \frac{\sum_{i=1}^{n} |V_i - \bar{V}_i|}{n}
\]

where \( V_i \) is measured volume for the ith tree, \( \bar{V}_i \) is predicted volume for the ith tree, \( \bar{V} \) is measured mean tree volume, \( n \) is the total number of trees.

**Results and discussion**

**Estimation and validation of volume equations**

The volume equations were classified as two types of functions: volume = f(DBH) with an independent variable

![Figure 1. Study area to develop stem volume model for Pinus densiflora, Pinus koraiensis, and Larix kaempferi in South Korea.](image)
and volume = f(DBH, height) with two independent variables. There are three equations in type 1 and five equations in type 2, and the parameter estimates and fit statistics were computed by species and all equations (Table 3). For the volume equations of type 1 with DBH, the parameter estimates were valid ($P < 0.05$), but the $P$-value of parameter b was not valid with $P = 0.7060$ in Equation (2) for $P. densiflora$. For this reason, Equation (2) failed to fit the model. Equation (3) had higher $R^2$ and lower RMSE and MAD than Equation (1). Especially, RMSE of $P. koraiensis$ was high with 0.0874 in Equation (3), but low with 0.0774 in Equation (7). Consequently, Equation (3) was evaluated as the most suitable one-variable equation with DBH.

For the volume equations of type 2 with DBH and height, parameter estimates of Equations (4) to (8) were firstly evaluated, and all parameters were valid ($P < 0.05$). Compared to other equations, however, Equations (5) and (8), especially for $L. kaempferi$, had fit statistics with low $R^2$ and high RMSE and MAD. Meanwhile, Equation (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 (5), (6), and (8) failed to be the best fit equations in this study.

Equations (4) and (7) had valid coefficients and good fit statistics for the three species. In case of Equation (7), however, the conversion of coefficients occurred the difference due to differences of assumption about the error term, and this logarithm form could lead to transformation bias (Burkhart and Tome 2012). Also, skewness and kurtosis of residuals by species showed Equation (4) was closer to normal distribution and more adequate than Equation (7), as in a previous study (Jeon et al. 2007). Consequently, Equation (4) was considered as the best model to predict the standing volume. This result corresponded to previous studies, which concluded that Equation (4), the so-called combined-variable function, is the best fit model (Burkhart 1977; Amateis and Burkhart 1987; Sherrill et al. 2011).

### Table 3. Coefficient and fit statistics of stem volume equations by species.

| Species       | Equation | Parameters | Fit statistics |
|---------------|----------|------------|---------------|
|               | $a$      | $b$        | $c$           | $d$   | $R^2$ | RMSE | MD  | MAD  |
| *Pinus densiflora* | 1        | -0.11515   | 0.00078265   | 0.000881 | 0.9171 | 0.1106 | 0.0000 | 0.0789 |
|               | 2        | -0.01972   | -0.00631     |         | 0.9174 | 0.1103 | 0.0001 | 0.0779 |
|               | 3        | -0.00577   | 0.0009042    |         | 0.9174 | 0.1104 | -0.0001 | 0.0780 |
|               | 4        | 0.0228     | 0.00003506   |         | 0.9729 | 0.0633 | 0.0004 | 0.0441 |
|               | 5        | 0.00003596 |           |         | 0.9720 | 0.0643 | 0.0061 | 0.0463 |
|               | 6        | 0.0405     | 0.000024     | 2.0277  | 1.0887 | 0.9730 | 0.0631 | -0.0036 | 0.0433 |
|               | 7        | -9.63696   | 1.79609      | 1.9148  |       | 0.9702 | 0.0560 | 0.0047 | 0.0443 |
|               | 8        | 54.1759    | 26732.4      |         |         | 0.9721 | 0.0642 | 0.0048 | 0.0463 |
| *Pinus koraiensis* | 1        | -0.23696   | 0.00098145   |         |         | 0.9661 | 0.0874 | 0.0000 | 0.0694 |
|               | 2        | 0.28478    | -0.03654     | 0.00158 |         | 0.9764 | 0.0729 | -0.0015 | 0.0546 |
|               | 3        | -0.01741   | 0.001234     |         |         | 0.9734 | 0.0774 | -0.0018 | 0.0600 |
|               | 4        | 0.01097    | 0.00003772   |         |         | 0.9861 | 0.0559 | 0.0001 | 0.0373 |
|               | 5        | 0.00003813 |           |         |         | 0.9860 | 0.0562 | 0.0036 | 0.0376 |
|               | 6        | -0.00042   | 0.000046     | 2.1767  | 0.7312  | 0.9872 | 0.0537 | -0.0056 | 0.0384 |
|               | 7        | -9.96556   | 2.01004      | 0.91845 |         | 0.9866 | 0.0549 | 0.0020 | 0.0380 |
|               | 8        | 129.8      | 23439.6      |         |         | 0.9867 | 0.0548 | -0.0022 | 0.0381 |
| *Larix kaempferi* | 1        | -0.09204   | 0.001016     |         |         | 0.9651 | 0.0964 | -0.0025 | 0.0767 |
|               | 2        | -0.47931   | 0.02622      | 0.000598|         | 0.9686 | 0.0914 | 0.0000 | 0.0775 |
|               | 3        | 0.0055     | 0.000108     |         |         | 0.9634 | 0.0988 | -0.0002 | 0.0786 |
|               | 4        | 0.05489    | 0.00003366   |         |         | 0.9934 | 0.0419 | 0.0000 | 0.0335 |
|               | 5        | 0.00003538 |           |         |         | 0.9894 | 0.0532 | 0.0019 | 0.0414 |
|               | 6        | -0.0277    | 0.000079     | 1.7176  | 1.0756  | 0.9965 | 0.0307 | -0.0034 | 0.0233 |
|               | 7        | -10.10382  | 1.78147      | 1.19935 |         | 0.9959 | 0.0332 | -0.0016 | 0.0248 |
|               | 8        | 186.2      | 23273.8      |         |         | 0.9905 | 0.0504 | 0.0119 | 0.0372 |

### Comparing volume equations

To check the developed volume equations of this study, scatterplots by species were displayed with measured volumes versus predicted volumes using Equation (4) (Figure 2). In addition, the volumes were compared with the volumes predicted by the coefficients of the previous studies in which Equation (4) had been used. That is, DBH and height of sample trees in this study were used for the predicted volume, and the measured volumes by legends were the same for all three species. The overall bias in the three species was minimized in predicted volumes used by the parameter estimates of this study, compared to other studies. The deviation between measured and predicted volumes were not significantly different in two-tailed paired $t$-test ($P = 0.9892$ for $P. densiflora$, $P = 0.9951$ for $P. koraiensis$, and $P = 0.9941$ for $L. kaempferi$).

The accuracy of predicted volumes varied by the species and parameters of previous studies. First, the predicted volumes for $P. densiflora$ started to be distinguished from true value as the measured volumes increased in Kim et al. (1994) and Korea Forest Research Institute (2012) (KFRI hereafter) and the measured and predicted volumes were significantly different ($P < 0.0000$) in both studies. However, the predicted volumes by Kim et al. (2001) were considerably similar to this study, and were not significantly different from measured volumes ($P = 0.0655$). It is judged that this characteristic of similarity or dissimilarity between this study and the previous studies was attributed to conformity of the study area and DBH and height distribution of the sample trees. Especially, the study area of Kim et al. (2001) was covered by the study area of this study, and the sample trees between Kim et al. (2001) and this study were comparable in terms of DBH and height.

In $P. koraiensis$, the predicted volumes in KFRI (2012) were not much different from the measured volumes even though the predicted were underestimated as volumes increased. However, there was a significant difference between the measured and predicted volumes of KFRI (2012).
in two tailed paired t-test (P = 0.0015). A large difference was shown between the volumes predicted by Shin (1994) and those measured (P < 0.0000). Shin (1994) studied small samples of trees of 14.8 cm in mean DBH and 9.9 m in mean height. We consider that this led to overestimation when the parameter estimates of Shin (1994) were applied to this study. In L. kaempferi, parameters of two previous studies predicted well, and the bias was smallest compared to P. densiflora and P. koraiensis (Shin et al. 1996; KFRI 2012). However, according to t-test, the measured and predicted volumes were significantly different in both previous studies (P < 0.0038 and P < 0.0001). Consequently, the parameter estimates in this study were the most accurate compared with other previous studies.

Predicted volumes using two-variable volume equations were displayed over DBH2H by species in order to work out volume over DBH and height and to compare the precision with other studies (Figure 3). The predicted volumes were overestimated or underestimated depending on the previous studies. Generally, Kim et al. (2001) for P. densiflora and Shin et al. (1996) for L. kaempferi were similar to this study. The volume model of this study is highly recommended for P. koraiensis, especially, in which DBH of a standing tree is bigger than small class (DBH < 16 cm) in South Korea.

Figure 2. Scatterplots for predicted versus measured volume by species comparing with previous studies through the best model of this study (V = a + bD2H).

Figure 3. Comparing predicted volume (V = a + bD2H) by species with previous studies.
**Best model selection and application**

In this study, Equations (3) and (4) were evaluated as the best model considering the validation of parameter estimates and fit statistics. However, foresters should decide a volume equation depending on the measured variables, only DBH or DBH and height, in the field. To compare Equations (3) and (4), volume deviation (predicted – measured) was presented as scatterplots by species (Figure 4). The deviation increased in both equations as the volume increased. Overall deviation was smaller in Equation (4) for all three species.

RMSE for *P. densiflora*, *P. koraiensis*, and *L. kaempferi* were 0.0633, 0.0559, and 0.0419 in Equation (4), and 0.1104, 0.0774, and 0.0988 in Equation (3). RMSE by species ranged from 0.0774 to 0.1104 in Equation (3) and from 0.0419 to 0.0633 in Equation (4). Also, other fit statistics including $R^2$ and MAD represented in Equation (4) were more accurate and precise. Consequently, Equation (4) was selected as the best model in this study. However, Equation (3) is considered to be applicable enough in the field if foresters can bear a little more deviation; MAD ranged from 0.06 to 0.0786 in Equation (3) and 0.0335 to 0.0441 in Equation (4). Not collecting height measurements to calculate the volume makes fieldwork more convenient and economic.

Finally, the volumes were presented over DBH to compare by species (Figure 5). The volumes were not much different until 20 cm in DBH, but it was gradually distinguished from each other as DBH increased. It is judged the different characteristics by species affects the volume growth such as height and taper. Thus, the volume based on DBH was *L. kaempferi*, *P. koraiensis*, and *P. densiflora* in sequence. This pattern was the same with the predicted volume using Equation (3), and so it proved the parameter estimates of this study are well predicted once more. However, the models of this study are highly cautioned to be applied within 16 cm to 50 cm in DBH in order to prevent extrapolation.

**Conclusion**

Data for *P. densiflora*, *P. koraiensis*, and *L. kaempferi* from Gangwon, Gyeonggi, North Gyeongsang provinces were used to develop the stem volume model. The best model was chosen among the eight stem volume equations widely used in forest biometrics. $V = a + bD^2H$ was the best model, considering the parameter validity and fit statistics. When compared with previous studies in Korea, this study’s model was found to be more accurate and applicable for the volume prediction of the study area.
In addition, two types of volume equations were compared to estimate the volume more effectively with time-saving fieldwork. \( V = a + bD^2H \) using DBH and height variables led to higher accuracy than \( V = aD + bD^2 \) using only DBH. However, the difference of MAD between the two models was 0.03, and if this is allowable, \( V = aD + bD^2 \) using DBH is evaluated to be applicable enough in the field considering the relative disadvantages of measuring the height variable. Finally, the volume was largest for \( L. kaempferi \) and smallest for \( P. densiflora \) comparing the stem volume of each species over DBH.

Consequently, the volume models of both \( V = a + bD^2H \) and \( V = aD + bD^2 \) show better suitability than the models of previous studies. Thus, these equations and parameter estimates are eminently practical for stem volume prediction of \( P. densiflora \), \( P. koraiensis \), and \( L. kaempferi \), especially in Gangwon, Gyeonggi, and North Gyeongsang provinces of South Korea.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This study was carried out with the support of ‘Science and Engineering Individual Basic Research Project (Project No. NRF-2016R1D1A1B02011648)’ funded by National Research Foundation of Korea in 2016.

**References**

Amateis RL, Burkhart HE. 1987. Tree volume and taper of loblolly pine varies by stand origin. South J Appl For. 11:185–189.

Avery TE, Burkhart HE. 2002. Forest measurements. 5th ed. New York (NY): McGraw-Hill.

Baskerville GL. 1972. Use of logarithmic regression in the estimation of plant biomass. Can J For Res. 2:49–53.

Beck DE. 1963. Cubic-foot volume tables for yellow-poplar in the Southern Appalachians. Research Note SE-16. Asheville (NC): US Forest Service, Southeastern Forest Experiment Station, Research Note SE-16: p. 4.

Burkhart HE. 1977. Cubic-foot volume of loblolly pine to any merchantable top limit. South J Appl For. 1:7–9.

Burkhart HE, Tomé M. 2012. Modeling forest trees and stands. New York (NY): Springer.

Clark III A, Saucier JR. 1990. Tables for estimating total-tree weights, stem weights, and volumes of planted and natural southern pines in the Southeast. Forest Research Paper 79. Dry Branch (GA): Georgia Forestry Commission; p. 23.

Honer TG. 1965. A new total cubic foot volume function. Forest Chron. 41:476–493.

Jeon BH, Lee SH, Lee YJ, Kim H, Kang HM. 2007. Estimation of site index and stem volume equations for \( Larix leptolepis \) stand in Jinan, Chonbuk. J Korean For Soc. 96:40–47. (in Korean with English abstract)

Kim KD, Choung SH. 1988. Equations for estimating volume and height of \( Larix leptolepis \) growing in the central region of Korea. J Korean For Soc. 77:23–31. (in Korean with English abstract)

Kim JS, Lee WK, Byun WH. 1994. Regional stem curve and volume function model of \( Pinus densiflora \) in Kangwon-province. J Korean For Soc. 83:521–530. (in Korean with English abstract)

Kim DG, Lee YJ, Lee DS, Hong SC. 2001. Estimation of stem volumes for \( Pinus densiflora \) for. ecta at Sokwangeri in Ulijin. Kor J For Meas. 4:1–7. (in Korean with English abstract)

Knoebel BR, Burkhart HE, Beck DE. 1984. Stem volume and taper functions for yellow-poplar in the Southern Appalachians. South J Appl For. 8:185–188.

Korea Forest Research Institute. 2012. Stem volume, biomass, and yield table. Seoul: Korea Forest Research Institute. (in Korean)

Korea Forest Research Institute. 2012a. Commercial tree species 1 \( Pinus densiflora \). Seoul: Korea Forest Research Institute. (in Korean)

Korea Forest Research Institute. 2012b. Commercial tree species 3 \( Pinus koraiensis \). Seoul: Korea Forest Research Institute. (in Korean)

Korea Forest Research Institute. 2012c. Commercial tree species 4 \( Larix kaempferi \). Seoul: Korea Forest Research Institute. (in Korean)

Korea Forest Service. 2016. 2016 Statistical yearbook of forestry. Daejeon: Korea Forest Service. (in Korean)

Lee D, Choi J. 2016. Estimating wood weight change on air drying times for three coniferous species of South Korea. J For Env Sci. 32:262–269.

Lee D, Choi J, Seo Y, Kim E. 2014. Nonlinear height-DBH growth models for \( Larix kaempferi \) in Gangwon and North Gyeongsang province. J For Env Sci. 30:201–207.

Myers CA. 1972. Volume, taper, and related tables for Southwestern Ponderosa pine (revised). Research Paper RM-2. Fort Collins (CO): US Forest Survey, Rocky Mountain Forest and Range Experiment Station; p. 25.

Park G, Lee D, Seo Y, Choi J. 2015. Height growth models for \( Pinus thunbergii \) in Jeju Island. J For Env Sci. 31:255–260.

Queen WR, Pienaar LV. 1977. Green and dry stem weight equations for site-prepared slash pine plantations in the coastal plain of Georgia and north Florida. Research Paper 1: Athens (GA): University of Georgia, School of Forest Resources, Plantation Management Research Co-Operative; p. 58.

SAS Institute Inc. 2013. SAS/STAT 9.4 user’s guide. Cary (NC): SAS Institute Inc.

Schumacher FX, Hall FDS. 1933. Logarithmic expression of timber-tree and stump biomass. Forest Soc. 65. (in Korean with English abstract)

Shin MY. 1994. A study on the development of individual tree volume equations for \( Pinus thunbergii \) in Gangwon-province. J Korean For Soc. 31:255–260.

Shin MY, Yun JW, Cha DS. 1996. Local correction of tree volume equation for \( Larix leptolepis \) by ratio-of-means estimator. J Korean For Soc. 85:65–66. (in Korean with English abstract)

Spurr SH. 1952. Forest inventory. New York (NY): Ronald Press Co.