

부트스트랩을 이용한 소나무의 목재기본밀도 추정 및 평가

표정기¹ · 손영모^{1*} · 김영환¹ · 김래현¹ · 이경학¹ · 이영진²

¹국립산림과학원, ²공주대학교

Use of a Bootstrap Method for Estimating Basic Wood Density for *Pinus densiflora* in Korea

Jung Kee Pyo¹, Yeong Mo Son^{1*}, Yeong Hwan Kim¹, Rae Hyun Kim¹,
Kyeong Hak Lee¹ and Young Jin Lee²

¹Division of Forest management, Korea Forest Research Institute, Seoul 130-712, Korea

²Department of Forest Resources, Kongju National University, Yesan 340-802, Korea

요약: 본 연구의 목적은 부트스트랩 시뮬레이션(Bootstrap simulation)을 이용하여 소나무의 목재기본밀도를 평가하고자 하였다. 소나무의 목재기본밀도는 생태형에 따라 강원지방소나무와 중부지방소나무의 자료로 구분하여 분석하였다. 비모수통계 방법의 하나인 부트스트랩 시뮬레이션 기법을 이용하여 추정된 목재기본밀도는 강원지방소나무에서 0.418(g/cm³), 중부지방소나무에서 0.464(g/cm³)으로 나타났다. 부트스트랩 시뮬레이션에서 100, 500, 1,000, 5,000번 반복 시행한 결과에 의하면, 모수 추정치의 95%신뢰구간은 일정한 수치로 나타난 반면에, 표준오차는 감소하는 경향으로 나타났다. 본 연구 결과로 제시된 목재기본밀도 추정치는 기존의 계수에 대한 단점을 보완하고, 신뢰성 높은 목재기본밀도 추정치로 적용이 가능할 것으로 사료된다.

Abstract: The purpose of this study was to develop the basic wood density (Abbreviated BWD) for *Pinus densiflora* and to evaluate the applicability of bootstrap simulation method. The data sets were divided into two groups based on eco-types in Korea, one from Gangwon type and the other from Jungbu type. The estimated BWDs derived from bootstrap simulation, which is one of the non-parametric statistics, were 0.418 (g/cm³) in the *Pinus densiflora* in Gangwon while 0.464 (g/cm³) in the *Pinus densiflora* in Jungbu. To evaluate the bootstrap simulation, the mean BWD, standard error and 95% confidence interval of probability density were estimated. The number of replication were 100, 500, 1,000, and 5,000 times that showed constant 95% confidence interval, while tended to decrease in terms of standard errors. The results of this study could be very useful to apply basic wood density values to calculate reliable carbon stocks for *Pinus densiflora* in Korea.

Key words : *Pinus densiflora*, basic wood density, bootstrap simulation, uncertainty, resampling

서론

지구 온난화를 방지하고 배출되는 온실가스의 감축을 위하여 국제적인 협약이 체결되고 구체적인 방안이 연구 중이다. 현재 진행 중인 기후 변화 협약에서 부속서 I(Annex I)국가는 온실가스의 배출과 흡수에 관한 통계 및 변화량에 관한 보고서를 국제적인 사무국에 제출하여야 한다. 국가보고서의 작성을 위하여 대기 중으로 배출되는 가스의 구분과 배출량에 관한 정도(精度) 높은 측정이 필요하다(Monte *et al.*, 1996; Kangas *et al.*, 2004). 산림과

직접적으로 관련되는 배출가스로 탄소가 있으며 탄소량을 파악하기 위하여 산림생태계의 다양한 부분에서 배출 계수가 연구되고 있다(Lehtonen *et al.*, 2004; IPCC, 2006). 임목은 산림의 주요한 탄소저장고이며, 이와 관련되는 배출계수는 목재기본밀도(Basic wood density), 바이오매스 확장계수(Biomass expansion factor), 뿌리함량비(Root-shoot ratio)가 있다. 현재, 지속적인 자료의 확충을 통해 수종별, 지역별 정도 높은 배출계수가 개발되고 있다(국립산림과학원, 2010).

기존의 배출계수 연구는 시간적, 경제적인 여건으로 인하여 자료 수집이 어려우며 계수 개발은 수집된 자료의 평균값으로 계산된다. 현재의 방법은 전국의 산림을 모집

*Corresponding author
E-mail: treelove@forest.go.kr

단으로 고려하며 표본의 선정, 자료의 수집, 배출계수 분석단계에 의해 다양한 이상치(Outlier)가 발생한다. 그리고 수집된 표본의 확률밀도(Probability density)가 정규성에 위배되거나 자료의 수가 적은 경우, 올바른 배출계수의 제시에 문제가 발생한다. 이러한 경우, 적용 가능한 기법이 비모수통계학(Non-parametric statistics)이다(Chemick, 2007). 부트스트랩 모형(Bootstrap simulation)은 비모수 통계의 일종으로 반복적인 표본추출을 통하여 통계량을 추정하거나 모집단의 경향을 파악하는데 이용된다(Gehring, 2006; Wasserman, 2006). 부트스트랩 모형은 다양한 부분에서 응용되고 있다. 김우중 등(2009), 김경섭(2010)은 부트스트랩 모형을 이용하여 기존의 점추정값(Point estimate)에서 발생하는 단점을 보완할 수 있다고 하였다. 김태규 등(2010)은 부족한 자료의 극대화를 통하여 기존의 방법과 비교하였으며, 정석근 등(2008)은 부트스트랩을 이용하여 모델의 신뢰구간을 검정하였다. 산림 부분에서 부트스트랩 모형은 배출계수의 개발보다 모형의 검증에 주로 이용되었다. Coulombe *et al.*(2010)은 가분비나무(Black spruce)의 수확모델을 부트스트랩 모형을 이용하여 검정하였으며 Gehring(2006)은 부트스트랩 모형을 이용하여 하안(河岸) 산림을 평가하였다. Nicholas *et al.*(2010)은 모형을 개발하고 적용하는 과정에서 모수의 민감도를 줄이는 방법으로 부트스트랩 모형을 적용하였다.

본 연구는 IPCC(2006)에서 권장하고 있는 배출계수의 추정 방법을 연구하고 비모수적 통계의 필요성을 검토한 후, 부트스트랩 모형을 이용하여 신뢰성 있는 소나무의 목재기본밀도를 제시하고자 하였다.

재료 및 방법

1. 분석자료

전국의 소나무 임분을 대상으로 총 42개소의 표준지를 선정하고, 각 표준지로부터 5분씩 총 210분의 표본목을 별채하였다. 생중량과 바이오매스를 추정하기 위한 절차는 다음과 같다. 1) 표본목은 조사지의 직경급을 고려하여 선정하고 피압목, 폭목, 정단부 피해목 등은 표본목의 선정에서 제외하였다. 2) 선정된 표본목은 지상부 0.2 m의 절단면부터 연륜과 벌도목의 길이를 측정하고, 재적계산 및 바이오매스 환산을 위해 단판을 채취하였다. 3) 채취한 단판으로부터 재적과 생중량을 측정하고 수피와 목질부를 분리하여 건중량을 측정하였다. 4) 목재기본밀도는 다양한 방법으로 계산이 가능하나, 본 연구에서는 수피를 제외한 생체적(Green volume)에 대한 목질부 건중량(Dry weight)의 비율로 계산하였으며 임목의 재적 측정을 통해 건중량의 전환이 가능하다(국립산림과학원, 2007).

IPCC(2006)는 동일 수종에서 다양한 조건에 따른 배출

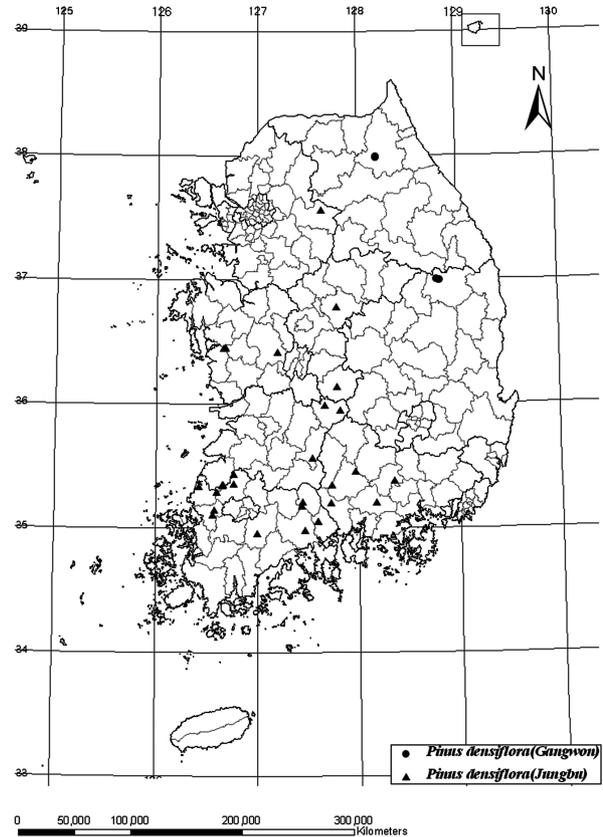


Figure 1. Geographical location of study sites for *Pinus densiflora*.

계수의 차이를 인정하고 계수를 구분하여 사용하는 것을 권장하고 있다. 따라서 본 연구에서는 소나무의 생태형에 따라 강원지방소나무와 중부지방소나무로 구분하였다(Figure 1). 그리고 IPCC가 정하는 기준입령(20년)별 목재기본밀도의 차이를 검정하기 위하여 t-검정을 실시하였다.

2. 부트스트랩 시뮬레이션

부트스트랩 모형은 비모수통계의 일종으로 표본수가 적거나 특별한 확률분포를 보이지 않을 경우에 사용하며 반복적으로 추출된 표본은 모집단의 분포를 따른다(Efron, 1979; Efron, 1981; Wasserman, 2006; Chernick, 2007). 부트스트랩 모형은 실제 자료에 근거하고 표본의 크기와 확률밀도함수의 편향에 크게 영향을 받지 않으며 표본과 유사한 통계량을 나타낸다(Chang *et al.*, 2009; Jeong and Kim, 2009). 본 연구에서 부트스트랩 모형의 적용절차는 다음과 같다. 1) 전체 표본($X=(x_1, x_2, x_3 \dots x_n)$)에서 무작위로 대표본을 추출(Random sampling)하여 부트스트랩 표본($X=(x_{1b}, x_{2b}, x_{3b} \dots x_{nb})$)을 생성하였다. 2) 부트스트랩 표본의 평균($\theta_b = \text{mean}(X_b)$)을 산정하였다. 3) 단계 1)과 2)를 반복하여 확률밀도함수(Probability density function)를 생성하였다. 4) 확률밀도함수에서 평균 및 95% 신뢰구간을 계산하였다. IPCC는 부트스트랩 표본의 반복추출횟수를

10,000번과 같이 사전에 설정하여 안정된 통계량을 구하도록 권장하고 있다. 본 연구는 시행횟수별 표본오차를 참고하여 확률밀도함수가 안정되는 부트스트랩의 시행횟수 (5,000번)를 설정하였다.

3. 부트스트랩 표본의 평가

부트스트랩 모형은 100, 500, 1,000, 5,000번마다 확률 밀도의 평균, 표본오차(SE_{boot}), 불확도(Uncertainty), 2.5%, 97.5%의 변화량을 계산하였다. 표본오차가 반복 횟수별 작아진다면 부트스트랩 모형을 통해 산정한 목재기 본밀도의 값은 의미가 있으며 95%구간의 경우, 정규성에 대한 가정 없이 추출된 목재기본밀도를 적은 값부터 순차 적으로 배열하여 실제 2.5%, 97.5%값을 산출한다(Efron, 1979; 전명식 등, 1997). 그리고 불확도는 확률밀도의 정 밀도와 변동성을 나타내므로(IPCC, 2006), 원표본의 불확 도에 비해 적은 값을 나타내거나 반복 횟수별 감소하는 경향을 보이면, 불확도의 산정은 유의미한 것으로 판단된 다. 부트스트랩 표본의 불확도는 확률밀도의 95%구간과 평균값의 비율로 계산되며 식 (1)과 같다(IPCC, 2006; Fujiwara, 2007).

$$Uncertainty = \frac{1/2 \times (95\%CI\ with)}{\mu} \times 100 \quad (1)$$

where, *CI* is confidence interval, μ is average of emission factor.

결과 및 고찰

1. 목재기본밀도의 개발

강원지방소나무는 강원도와 경상북도를 중심으로 12개 소의 표준지를 선정하고 총 60본의 표본목을 벌채하였다. 이외의 지역은 중부지방소나무로 분류하여 30개소의 표 준지에서 총 150본의 표본목을 벌채하였다. 수집된 표본 목의 영급은 IV영급으로 우리나라 소나무의 평균영급과 유사하였다. 강원지방소나무가 중부지방소나무에 비해 임 령은 낮지만 평균 흉고직경과 수고는 높게 나타났다(Table 1). 생태형에 따른 강원지방소나무와 중부지방소나무 두 집단 간 목재기본밀도의 차이를 검정하기 위하여 SAS (2006)를 이용하였으며 t-검정을 실시한 결과, 두 집단 간 에는 통계적인 유의성($P\text{-value} < 0.0001$)이 나타난 반면에,

Table 1. The summary of observed statistics for *Pinus densiflora*.

Eco-types	No. of observations	Age (years)	DBH (cm)	Height (m)
Gangwon	60	35.8	24.2	15.0
		9.0–82.0	5.8–59.4	4.6–24.9
Jungbu	150	37.4	17.0	11.8
		9.0–70.0	3.6–34.3	3.0–20.8

Table 2. Fit statistics and uncertainty of basic wood density for *Pinus densiflora*.

Eco-types	BWD (g/cm^3)	$SE_{\bar{x}}$	2.5 percentile	97.5 percentile	Uncertainty (%)
Gangwon	0.418	0.008	0.319	0.525	24.67
Jungbu	0.464	0.005	0.305	0.617	33.64

IPCC(2006)가 권장하는 20년 단위의 영급에 따른 목재기 본밀도는 강원지방소나무와 중부지방소나무에서 각각 통 계적 차이($P\text{-value}=0.1358, 0.1925$)는 나타나지 않았다. 따 라서, 소나무의 목재기본밀도의 경우 생태형에 따른 구분 만이 필요하며, 영급에 대해서는 동일한 계수의 적용이 가 능한 것으로 사료된다.

2. 목재기본밀도의 불확도 평가

목재기본밀도는 강원지방소나무에서 $0.418 (g/cm^3)$, 중 부지방소나무에서 $0.464 (g/cm^3)$ 으로 나타났다. 이러한 차 이는 강원지방소나무와 중부지방소나무의 건중량보다 재 적 차이에 기인하는 것으로 사료된다. 수집된 표본목은 각 지역의 환경적, 지역적 차이에 따라 직경과 수고성장 패 턴이 상이하여 목재기본밀도 산출에 영향을 준다는 기존 의 연구 결과와 유사한 것으로 나타났다(박인협 등, 2005). 강원지방소나무의 목재기본밀도에 대한 불확도는 24.67 (%), 중부지방소나무는 33.64 (%)로 나타났다(Table 2). 강 원지방소나무의 불확도가 낮게 나오는 것은 자료의 개수 보다 확률밀도의 95% 신뢰구간의 차이 때문인 것으로 사 료된다. 또한, 중부지방소나무에서 넓은 95% 신뢰구간을 나타내는 이유는 조사지역이 환경적으로 다양하고 넓은 지역에 분포하기 때문인 것으로 사료된다. FAO (2006)는 적용 가능한 불확도를 40% 이하로 제시하였으며, 본 연 구에서 계산된 불확도는 40% 이내로 나타났다. 그러나 Fujiwara 등(2007)이 연구한 소나무의 불확도 연구와 비교 하면 높게 나타났다. 이와 같은 차이의 원인은 표본의 선 정, 표본의 갯수, 자료의 수집, 목재기본밀도를 분석하는 단계에서 발생하는 불확도가 누적된 것으로 판단된다. 따 라서, 불확도를 줄이고 정도(precision) 높은 목재기본밀도 의 산정을 위해서는 많은 수의 표본을 수집하고 환경적, 지역적 변수에 대한 분석이 필요한 것으로 사료된다(Monte *et al.*, 1996; Macfarlane *et al.*, 2000; Refsgaard *et al.*, 2007).

3. 부트스트랩 모형(Bootstrap simulation) 추정치 비교

표본으로부터 산출된 목재기본밀도의 평균값과 부트스 트랩 시뮬레이션의 결과를 비교한 결과는 Table 3과 같다. 부트스트랩 모형의 시행횟수 증가에 따른 평균과 원표본 의 평균은 유사한 값으로 나타났다. 부트스트랩 표본의 표

Table 3. Bootstrap simulation results of basic wood density for *Pinus densiflora*.

Eco-types	Original data	No. of Rep.	Boot. sample	SE _{boot}	2.5 percentile	97.5 percentile	Uncertainty (%)
Gangwon	0.418	100	0.417	0.0008	0.401	0.432	3.714
		500	0.417	0.0003	0.402	0.434	3.790
		1000	0.418	0.0002	0.403	0.434	3.749
		5000	0.417	0.0001	0.402	0.433	3.730
Jungbu	0.464	100	0.464	0.0005	0.452	0.476	2.490
		500	0.464	0.0002	0.452	0.475	2.525
		1000	0.463	0.0001	0.452	0.475	2.449
		5000	0.464	0.0000	0.452	0.475	2.441

본오차는 반복추출 횟수의 증가에 따라 감소하는 경향으로 나타났으며 2.5%, 97.5%의 값은 일정하게 나타났다. 시행횟수 증가를 통해 부트스트랩 표본의 평균은 일정한 값으로 나타났다. 본 연구의 결과, 원표본의 확률밀도에 대한 기초 통계량을 검정하고 이상치가 발견되거나 자료의 수가 부족한 경우, 배출계수의 적용에는 부트스트랩 모형을 적용하여 평가하는 것이 바람직한 것으로 사료된다.

결론

우리나라의 대표수종인 소나무를 대상으로 부트스트랩 모형을 이용한 목재기본밀도의 산출과 이용의 타당성을 검정하였다. IPCC가 정하는 임령기준(20년)과 생태형에 따른 목재기본밀도의 차이를 검정을 위하여 t-검정을 실시한 결과, 생태형에 따른 목재기본밀도는 유의한 차이가 나타났지만 임령에 따른 차이는 없는 것으로 나타났다. 목재기본밀도의 평균은 강원지방소나무에서 0.418 (g/cm³), 중부지방소나무에서 0.464 (g/cm³)으로 나타났으며, 불확도는 강원지방소나무에서 24.67%, 중부지방소나무에서 33.64%로 나타났다.

불확도 평가는 비모수통계 방법인 부트스트랩 방법을 이용하여 기존의 목재기본밀도 추정방법과 부트스트랩 모형의 모의실험 결과를 비교분석 하였다. 부트스트랩 모형의 시행횟수를 100, 500, 1,000, 5,000번으로 설정하여 시행 횟수별 목재기본밀도를 산정하고 유용성을 검정하기 위하여 표본오차(SE_{boot})와 2.5%, 97.5%구간의 값을 비교하였다. 부트스트랩 시행횟수의 증가에 따라 목재기본밀도와 95%구간의 값은 일정하지만 표본오차는 감소하는 추세를 나타냈다. 본 연구에서 사용된 부트스트랩 모형은 현재 사용되고 있는 배출계수를 보완할 수 있으며 신뢰성 있는 배출계수를 제공하는데 도움이 될 것으로 사료된다.

감사의 글

본 연구는 산림청 ‘산림탄소계정 기반 구축 연구’ 조사 연구 결과의 일부임.

인용문헌

1. 국립산림과학원. 2007. 산림 바이오매스 및 토양탄소 조사·분석 표준. 국립산림과학원. pp. 4.
2. 국립산림과학원. 2010. 산림 온실가스 인벤토리를 위한 주요 수종별 탄소배출계수. 국립산림과학원. pp. 7.
3. 김경섭. 2010. Bootstrap 기법을 이용한 BOD 평균 농도 및 신뢰구간 분석. 수질보전 한국물환경학회지 26(2): 297-302.
4. 김우중, 강기훈. 2009. 붓스트랩을 이용한 다차원척도법의 효율성 연구. 한국데이터정보과학회지 20(2): 301-309.
5. 김태규, 박상규, 하명호. 2010. 안정성 연구에서의 사용기간에 관한 비모수적 추론. 품질경영학회지 39(1): 96-100.
6. 박인협, 박관수, 이경학, 손영모, 서정호, 손요환, 이영진. 2005. 소나무의 생태형과 임령에 따른 물질 현존량 확장계수. 한국임학회지 94: 441-445.
7. 정석근, 최일수, 장대수. 2008. 부트스트랩과 베이지안 방법으로 추정된 수산자원관리에서의 생물학적 기준점의 신뢰구간. 한국수자원학회지 41(2): 107-112.
8. 전명식, 정형철, 진서훈. 1997. 붓스트랩방법의 이해. 자유아카데미. pp. 3-9.
9. Chang, K.Y., Hong, K.O. and Park, S.I. 2007. Bootstrap simulation for quantification of uncertainty in risk assessment. Korean Journal of Veterinary Research 47(2): 259-263.
10. Chernick, M.R. 2007. Bootstrap Methods: A Guide for Practitioners and Researchers. John Wiley and Sons. 53-58.
11. Coulombe, S., Bernier, P.Y. and Raulier, F. 2010. Uncertainty in detecting climate change impact on the projected yield of black spruce (*Picea mariana*). Forest Ecology and Management 259: 730-738.
12. Efron, B. 1979. Bootstrap methods: Another look at the jackknife. The annals of statistics 7: 1-26.
13. Efron, B. 1981. Non parametric estimates of standard error: The jackknife, the bootstrap and other methods. Biometrika 68(3): 589-99.
14. FAO. 2006. Global forest resources assessment 2005. Food and Agriculture Organization, Rome, Italy.
15. Fujiwara, T., Kana, Y. and Kuroda, K. 2007. Basic den-

- sities as a parameter for estimating the amount of carbon removal by forests and their variation. 森林綜合研究所報告 6(4): 215-226.
16. Gehringer, K.R. 2006. Structure-based nonparametric target definition and assessment procedures with an application to riparian forest management. *Forest Ecology and Management* 223: 125-138.
 17. IPCC. 2006. 2006 IPCC Guidelines for National Greenhouse Gas Inventories. Volume 1. General Guidance and Reporting. IPCC National Greenhouse Gas Inventory Programme. Institute for Global Environmental Strategies. pp. 3.6-3.78.
 18. IPCC. 2006. 2006 IPCC Guidelines for National Greenhouse Gas Inventories. Volume 4. Agriculture, Forestry and Other Land Use. IPCC National Greenhouse Gas Inventory Programme. Institute for Global Environmental Strategies. pp. 4.73.
 19. Jeong, H.C. and Kim, D.H. 2009. Constructing simultaneous confidence intervals for the difference of proportions from multivariate binomial distributions. *The Korean Journal of Applied Statistics* 22(1): 129-140.
 20. Kangas, A.S. and Kangas, J. 2004. Probability, possibility and evidence : approaches to consider risk and uncertainty in forestry decision analysis. *Forest Policy and Economics* 6: 169-188.
 21. Lehtonen, A., Mäkipää, R., Heikkinen, J., Sievänen, R. and Liski, J. 2004. Biomass Expansion factor (BEF) for Scots pine, Norway spruce and birch according to stand age for boreal forests. *Forest Ecology and Management* 188: 211-224.
 22. Macfarlane, D.W., Edwin, J.G. and Harry, T.V. 2000. Incorporating uncertainty into the parameters of a forest process model. *Ecological Modelling* 134: 27-40.
 23. Monte, L., Lars, H., Ulla, B., John, B. and Rudie, H. 1996. Uncertainty analysis and validation of environmental models : The empirically based uncertainty analysis. *Ecological Modelling* 91: 139-152.
 24. Refsgaard, J.C., Jeroen, P.V.D.S., Anker, L.H. and Peter, A.V. 2007. Uncertainty in the environmental modelling process-A framework and guidance. *Environmental Modelling and Software* 22: 1543-1556.
 25. SAS Institute, Inc. 2006. SAS/STAT 9.1.3 User's Guide. SAS Institute, Inc. Cary. NC.
 26. Vaughn, N.R., Turnblom, E.C. and Ritchie, M.W. 2010. Bootstrap evaluation of a young douglas-fir height growth model for the pacific northwest. *Forest Science* 56(6): 592-602.
 27. Wasserman, L. 2006. All of nonparametric statistics. Springer. pp. 27-41.

(2011년 4월 12일 접수; 2011년 4월 22일 채택)