한국조경학회지 54(5): 156~174, 2026. 10.
J. KILA Vol. 54, No. 5, pp. 156~174, October, 2026
pISSN 1225-1755 eISSN 2288-9566
https://doi.org/10.9715/KILA.2026.54.5.156

# 낙엽활엽수의 수고-흉고직경 조건부 분포 분석을 위한 수종별 분위수 회귀식 구축†

## Development of Species-Specific Quantile Regression Equations for Analyzing the Conditional Height-Diameter Distribution of Deciduous Broadleaved Trees†

**최호준*, 이정은**, 조휘리**, 전진형*****
*고려대학교 환경생태공학과 석박사통합과정, **고려대학교 환경생태공학과 석사과정, ***고려대학교 환경생태공학부 교수

**Choi, Hojun*, Lee, Jeongeun**, Jo, Hwiri**, Chon, Jinhyung*****
*Integrated M.S. and Ph.D. Student, Dept. of Environmental Science and Ecological Engineering, Korea University
**Master's Student, Dept. of Environmental Science and Ecological Engineering, Korea University
***Professor, Division of Environmental Science and Ecological Engineering, Korea University

**Received:** July 1, 2026
**Revised:** July 29, 2026 (1st)
September 7, 2026 (2nd)
**Accepted:** September 9, 2026
3인익명 심사필

**Corresponding author :**
Jinhyung Chon
Professor, Division of Environmental Science and Ecological Engineering, Korea University, Seoul 02841, Korea
Tel.: +82-2-3290-3621
E-mail: jchon@korea.ac.kr

†본 논문은 고려대학교 OJERI(Supported by OJEong Resilience Institute, Korea University) 연구비에 의하여 수행되었음.

**국문초록**

본 연구의 목적은 낙엽활엽수종별 수고-흉고직경 관계의 조건부 하위 및 상위 분위범위를 추정하는 분위수 회귀식을 구축하고, 그 예측성능과 회귀선의 안정성 및 신뢰성을 평가하는 데 있다. 분석에는 2006-2020년 제5-7차 국가산림자원조사 자료를 활용하였으며, 온대중부 적용 가능 낙엽활엽교목 및 아교목 가운데 유효 관측값 100개 이상과 고유 표본점 30개 이상의 기준을 충족한 25개 수종, 64,072개 관측값을 최종 분석대상으로 선정하였다. 수종별 q10 및 q90 분위회귀에서 선형, 평균중심화 이차 및 cubic B-spline 모형을 구축하고, 표본점 단위 5-fold 교차검증의 pinball loss와 1 표준오차 규칙을 이용하여 최종모형을 선정하였다. 분석 결과, 총 25개의 수종 중 가죽나무와 갈참나무에서는 이차모형이, 나머지 23개 수종에서는 선형모형이 선정되었다. q10과 q90 관측비율의 수종 평균은 각각 0.105와 0.896이었으며, q10-q90 구간의 중앙 80% 포괄률은 0.791로 나타났다. 관측된 실측수고 범위에서 분위회귀선의 교차는 확인되지 않았으며, 1,000회 표본점 단위 군집 부트스트랩을 통해 회귀계수의 95% 신뢰구간을 산정하였다. 본 연구에서 구축한 회귀식은 개별 수목의 미래 생장이나 도시수목의 생육상태를 판정하는 기준이 아니라, 동일한 실측수고 조건에서 산림수목의 흉고직경 조건부 분포를 비교하기 위한 탐색적 비교자료이다. 향후 도시숲과 가로수의 독립적인 실측자료를 활용한 외부검증이 필요할 것으로 사료된다.

**주제어:** 조건부 분위범위, 국가산림자원조사, 현장측정자료, 군집 부트스트랩, 실측수고

**ABSTRACT**

This study aimed to develop species-specific quantile regression equations for estimating the lower and upper conditional quantile ranges of the height-diameter relationship in deciduous broadleaved trees and to evaluate their predictive performance and the stability and reliability of the regression curves. Data from the fifth to seventh Korean National Forest Inventory (2006-2020) were used for analysis. Among temperate central deciduous broadleaved canopy and understory tree species, 25 species and 64,072 observations that satisfied the criteria of at least 100 valid observations and 30 unique sample plots were selected for final analysis. For each species, linear, mean-centered quadratic, and cubic B-spline models were fitted for the q10 and q90 quantile regressions, and the final model was selected using plot-level five-fold cross-validation based on pinball loss and the one-standard-error rule. As a result, quadratic models were selected for *Ailanthus altissima* and *Quercus aliena*, whereas linear models were selected for the remaining 23 species. The species-level mean observed proportions for q10 and q90 were 0.105 and 0.896, respectively, and the coverage of the central 80% interval between q10 and q90 was 0.791. No crossing of the quantile regression curves was observed within the range of measured tree heights, and 95% confidence intervals for the regression coefficients were estimated using 1,000 plot cluster bootstrap replications. The equations developed in this study are not intended to serve as criteria for assessing future growth of individual trees or the growth condition of urban trees, but rather as exploratory reference equations for comparing the conditional DBH distribution of forest trees under the same measured height condition. External validation using independent measured data from urban forests and street trees will be needed in future research.

**Keywords:** Conditional Quantile Range, National Forest Inventory, Field-Measured Data, Cluster Bootstrap, Diameter at Breast Height

# 1. 서론

도시숲은 그늘과 증발산을 통한 미기후 조절, 탄소 흡수, 생물다양성 보전과 같은 환경적 기능을 제공하며(Amini Parsa et al., 2019; Rahim et al., 2024), 휴양 및 경관 제공, 신체적·정신적 건강 증진 등의 사회경제적 혜택을 함께 제공함으로써(Grima et al., 2020; Pirselimoğlu Batman and Ender Altay, 2025) 건강하고 회복력 있는 도시를 조성한다(Almas and Conway, 2016).

그러나 이처럼 다양한 혜택을 제공하는 도시숲은 최근 높은 온도와 가뭄, 강우 패턴 변화와 같은 기후변화의 영향을 직접적으로 받고 있으며, 특히 폭염과 가뭄으로 인한 열 스트레스가 도시숲의 수목 생장 저하를 가속화하고 있다(Allen et al., 2015; Marchin et al., 2025).

이러한 생육환경 저하는 도시에서 더욱 복합적으로 나타난다. 최근 연구들은 기후변화로 인해 전 세계 도시 수목의 상당수가 기존의 적정 생육환경을 벗어난 고온 및 수분 스트레스에 노출될 가능성이 증가하고 있으며, 고온 및 수분 스트레스에 따른 생장 저하 위험도 증가하고 있음을 보고하고 있다(Esperon-Rodriguez et al., 2022; Huang, 2022). 또한 도시 환경에서는 포장면 증가, 불투수면 확대 등에 따른 수분 공급 제한, 건물과 도로로 인한 열섬 현상, 제한된 식재 공간 등 다양한 인위적 요인이 수목 생장에 추가적인 스트레스로 작용하고 있다(Pretzsch et al., 2017). 이처럼 다양한 스트레스 요인이 존재하는 이질적인 도시 환경에서 도시숲의 수목 생장 저하는 관리에 대한 경제적 부담을 초래할 뿐만 아니라 수목이 쇠퇴함에 따라 도시숲이 제공하는 다양한 혜택이 감소하는 결과를 야기할 수 있다.

따라서 도시숲의 양적 확대가 생태적, 사회경제적 혜택의 지속성으로 연결되기 위해서는 조성 이후 수목의 크기와 구조적 특성을 정기적으로 기록하고, 수종과 크기 조건에 따라 비교할 수 있는 관리체계를 마련할 필요가 있으며, 이러한 관리체계가 실효성을 갖기 위해서는 수종별, 크기별 차이를 반영할 수 있는 정량적 비교 기준이 필요하다.

이러한 상대적 분포 기준을 설정하기 위해서는 수목의 구조적 특성을 설명할 수 있는 지표를 활용하는 것이 필요하다. 특히 흉고직경과 수고는 수목의 구조와 안정성, 생장 상태를 이해할 수 있는 지표이며(Kang et al., 2021; Ige and Komolafe, 2022), 수목의 크기 및 성장 추정과 생장 특성 분석에 활용되고 있는 대표적인 지표이다(Sumida et al., 2013; Scarmana and Mcdougall, 2023; Muscas et al., 2023). 그 외에도 수고와 흉고직경 지표는 비율을 통해 수목의 활력도를 평가하기도 하며, 수목의 성장 추정에도 활용될 수 있다(Sharma et al., 2019; Zhang et al., 2020). 수고와 흉고직경은 현장에서 반복적으로 측정할 수 있는 대표적인 구조변수이므로, 동일 수종에서 수고가 비슷한 수목들의 흉고직경 분포를 파악하면 개별 수목의 흉고직경이 다른 수목에 비해 상대적으로 가는지 또는 굵은지를 비교할 수 있다.

기존의 수고-흉고직경 연구는 주로 흉고직경을 이용하여 평균 수고 또는 수고-흉고직경비를 예측하는 회귀모형을 구축하거나 수종, 경쟁상태, 입지조건에 따른 관계의 차이를 설명하는 데 초점을 두어 왔다(Sharma et al., 2019; Zhang et al., 2020). 이러한 평균 중심의 모형은 수고와 흉고직경 사이의 전반적인 관계를 설명하는 데 유용하지만, 동일한 수고 조건에서 흉고직경이 형성하는 하위 및 상위 분포의 차이를 하나의 평균선만으로 제시하기는 어렵다.

따라서 흉고직경의 절댓값이나 평균값만을 제시하는 것에서 나아가, 동일한 수고 조건에서 흉고직경이 형성하는 조건부 분포를 파악할 필요가 있다. 이러한 조건부 분포는 개별 수목의 흉고직경이 같은 수고를 가진 다른 개체들과 비교하여 상대적으로 작은지, 중앙 범위에 속하는지 또는 상대적으로 큰지를 정량적으로 표현할 수 있게 한다. 분위회귀는 조건부 평균이 아니라 분포의 특정 위치를 추정하므로, 동일한 실측수고 조건에서 상대적으로 작은 흉고직경과 큰 흉고직경이 각각 어떠한 관계를 나타내는지를 구분하여 제시할 수 있다.

다만 도시 수목의 수고와 흉고직경의 관계를 분석하기 위해서는 수종별 개체목 단위의 충분한 현장자료가 필요하지만, 현재 국내 도시숲을 대상으로 장기간 축적된 표준화된 자료는 제한적이다. 따라서 본 연구는 전국 단위에서 수집된 국가산림자원조사 자료를 활용하여 수종별 수고-흉고직경 관계의 기초 분포 범위를 우선적으로 검토하고자 하였다.

이를 바탕으로 본 연구의 목적은 국가산림자원조사 자료를 활용하여 낙엽활엽수종별 실측수고-흉고직경 관계의 조건부 10% 및 90% 분위회귀모형을 구축하고, 동일한 실측수고 조건에서 나타나는 흉고직경의 하위 및 상위 분위

범위를 산림수목 기준의 탐색적 비교자료로 제시하는 데 있다. 구체적으로는 첫째, 수종별 선형, 이차, cubic B-spline 후보모형의 표본점 단위 검증성능을 비교하여 최종 함수형태를 선정하고, 둘째, 최종모형의 q10 및 q90 관측비율과 중앙 80% 구간 포괄률, 분위회귀선의 교차 여부 및 회귀계수의 불확실성을 평가하며, 셋째, 최종 분석 대상 전체 수종의 q10 및 q90 회귀식과 실측수고 범위를 제시하고자 하였다. 본 연구의 결과는 도시수목의 생육상태를 직접 판정하거나 개별 수목의 미래 생장을 예측하는 기준이 아니라, 도시수목 실측자료가 충분히 축적되지 않은 상황에서 산림수목의 구조적 분포와 비교하고 후속 현장조사 및 외부검증을 설계하기 위한 기초자료로 활용하고자 한다.

## 2. 연구 방법

### 2.1 분석자료 및 수종 선정

본 연구에서는 수종별 수고-흉고직경 관계의 조건부 분포를 분석하기 위해 국가산림자원조사(National Forest Inventory, NFI) 자료를 활용하였다(산림청, 2022a; 2022b; 2022c). NFI 자료는 전국 산림을 대상으로 표준화된 조사체계에 따라 구축된 대규모 자료로서, 다양한 지역과 입지 및 임분조건에 분포하는 수목의 구조적 특성을 수종별로 비교하는 데 활용될 수 있다(강진택, 2019; Sharma et al., 2019). 다만 NFI 자료는 도시숲이나 가로수를 대상으로 수집된 자료가 아니므로, 본 연구에서는 분석 결과를 도시수목의 생육상태를 직접 판정하는 기준이 아니라 산림수목의 수고-흉고직경 관계에 관한 탐색적 비교자료로 한정하였다.

분석대상 기간은 제5차, 제6차 및 제7차 국가산림자원조사에 해당하는 2006년부터 2020년까지로 설정하였다. 연도별 NFI 원자료의 임목조사표에서 집락번호, 표본점번호, 조사차기, 조사연도, 개체목 번호, 학명번호, 수종명, 흉고직경, 실측수고, 개체목 구분코드 및 수령을 추출하였다.

NFI 임목조사표에는 현장에서 기록된 실측수고와 수고곡선식으로 산정된 추정수고가 함께 수록되어 있다. 추정수고는 흉고직경을 포함한 수고곡선식으로 산정되므로, 이를 다시 흉고직경의 설명변수로 사용할 경우 두 변수 사이에 순환논리가 발생할 수 있다. 이에 본 연구에서는 추정수고를 분석에서 제외하고, 현장에서 측정된 실측수고와 흉고직경이 함께 기록된 관측값만을 사용하였다. 실측수고는 모든 조사 개체목에 기록된 변수가 아니므로, 본 연구의 분석대상은 NFI 전체 조사목이 아니라 실측수고와 흉고직경이 모두 기록된 개체목으로 한정하였다.

수종 후보군은 2025년도 도시숲경관 사업계획에 제시된 온대중부 적용 가능 수종을 기준으로 설정하였다. 해당 계획은 국내 도시숲 적용 가능 수종을 온대중부, 온대남부 및 난온대 유형으로 구분하여 제시하고 있다. 본 연구에서는 이 가운데 온대중부 수종군에 포함된 낙엽활엽교목 및 낙엽활엽아교목 52종을 1차 후보수종으로 선정하였다(표 1 참조). 침엽교목은 낙엽활엽수와 수관구조, 줄기 형태 및 수고-흉고직경 관계가 다르게 나타날 수 있으므로, 생육형이 비교적 유사한 수종군 내에서 결과를 비교하기 위해 후보수종에서 제외하였다.

분석자료를 구축하기 위해 표 1에 제시된 52개 1차 후보수종과 NFI 원자료의 수종명을 대조한 후, 후보수종에 해당하는 개체목 자료만을 추출하였다. 이에 따라 NFI 원자료에 존재하더라도 표 1의 후보수종 목록에 포함되지 않은 수종은 분석자료 구축 단계에서 제외하였다. 추출된 자료 가운데 실측수고 또는 흉고직경이 결측이거나 0 이하인 관측값을 제외하였으며 수종별 실측수고와 흉고직경의 최솟값과 최댓값을 검토하여 입력 오류 또는 생물학적으로 해석하기 어려운 극단값의 존재 여부를 확인하였다.

이후 수종별 분위회귀모형의 추정에 필요한 관측값과 표본점 단위 검증에 필요한 집단 수를 확보하기 위해 유효 관측값이 100개 이상이고 고유 표본점이 30개 이상인 수종을 주 분석대상으로 선정하였다. 유효 관측값 100개 기준은 절대적인 통계 임계값이라기보다, q10 및 q90과 같은 꼬리 분위수를 추정할 때 지나치게 작은 표본으로 인해 회귀선이 불안정하게 추정되는 문제를 줄이기 위한 기준으로 적용하였다. 또한 고유 표본점 30개 기준은 특정 수종

표 1. 1차 후보수종 52종

| 구분 | 주요 수종 |
|---|---|
| 온대 중부 교목, 아교목(낙엽 활엽수) | 가래나무, 가죽나무, 갈참나무, 감나무, 겹벚나무, 계수나무, 고로쇠나무, 개오동, 꽃산딸나무, 느티나무, 능수버들, 단풍나무, 당단풍나무, 대왕참나무, 대추나무, 마가목, 매화나무, 모감주나무, 모과나무, 목련, 무궁화, 박태기나무, 밤나무, 백합나무, 버드나무, 벚나무, 벽오동, 복자기, 뽕나무, 산딸나무, 산사나무, 산수유, 살구나무, 상수리나무, 양버들, 양버즘나무, 왕벚나무, 은사시나무, 이팝나무, 자귀나무, 자두나무, 자작나무, 중국단풍, 쪽동백나무, 층층나무, 칠엽수, 팥배나무, 팽나무, 피나무, 함박꽃나무, 황벽나무, 회화나무 |

의 관측값이 소수의 표본점에 편중되는 상황을 완화하고, 표본점 단위 5-fold 교차검증과 군집 부트스트랩을 수행할 수 있을 만큼의 집단 수를 확보하기 위해 설정하였다. 이러한 자료 선별 절차를 거쳐 유효 관측값 수와 고유 표본점 수 기준을 모두 충족한 수종을 최종 분석대상으로 선정하였다.

## 2.2 분석도구 및 활용 Python 라이브러리

모든 자료처리와 통계분석은 Python 3.9.6 환경에서 수행하였다. 국가산림자원조사 원자료의 불러오기, 결측자료 처리, 변수 변환, 수종별 자료 분할, 반복 표본점 및 반복 개체 후보 확인, 기초통계량 산정과 결과표 작성에는 pandas 2.3.1과 NumPy 2.0.2를 활용하였다.

수종별 q10 및 q90 분위회귀모형은 statsmodels 0.14.6의 QuantReg를 이용하여 추정하였다. 수종별 선형, 이차 및 cubic B-spline 후보모형의 예측성능을 비교하기 위한 표본점 단위 5-fold 교차검증에는 scikit-learn 1.6.1의 GroupKFold를 사용하였으며, cubic B-spline 기저함수의 생성에는 동일 라이브러리의 SplineTransformer를 활용하였다. 표본점 단위 군집 부트스트랩의 난수 생성과 복원추출에는 NumPy의 default_rng를 사용하였으며, 분석의 재현성을 확보하기 위해 고정된 난수 시드를 적용하였다.

Cubic B-spline 후보모형에서는 SplineTransformer로 생성한 기저변수를 각 훈련 fold에서 표준화하고, 동일한 변환을 해당 검증 fold에 적용하였다. 이 절차는 기저변수의 척도 차이로 인한 수치적 불안정성을 줄이면서 검증자료의 정보가 훈련과정에 유입되는 것을 방지하기 위해 적용하였다.

## 2.3 분위회귀모형 설정 및 평가 방법

본 연구는 실측수고가 주어졌을 때 흉고직경 조건부 분포의 특정 위치를 추정하기 위해 분위회귀분석을 적용하였다. 조건부 평균을 추정하는 최소제곱회귀와 달리 분위회귀는 q10, q50, q90과 같이 관심 있는 조건부 분위함수를 직접 추정하므로, 분포가 비대칭이거나 수고 구간에 따라 산포가 달라지는 자료에서 하위와 상위 관계를 구분할 수 있다(Koenker, 2017; Zhang et al., 2020).

본 연구에서는 흉고직경을 종속변수로, 실측수고를 설명변수로 설정하였다. 회귀계수의 해석 편의성과 비선형모형의 수치적 안정성을 고려하여 실측수고는 원자료에 기록된 cm 단위 값을 m 단위로 환산하였으며, 흉고직경은 원자료의 cm 단위를 유지하였다. 본 연구에서 추정한 모형은 동일한 개체목의 시간에 따른 변화를 분석하는 생장모형이 아니라, 2006-2020년 여러 조사차수를 통합한 자료에서 시간적 변화가 아닌 개체목 간 실측수고와 흉고직경의 구조적 관계를 분석하는 모형이다. 따라서 모형의 계수와 예측값은 개별 수목의 미래 생장을 의미하는 것이 아니라, 통합된 조사자료에서 동일한 실측수고 조건에 나타나는 흉고직경의 조건부 분포로 해석하였다.

흉고직경 조건부 분포의 하위 및 상위 위치를 함께 파악하기 위해 q10과 q90을 분석 분위로 설정하였다. q10은 동일한 실측수고 조건에서 흉고직경의 조건부 10% 분위값으로, 관측값의 약 10%가 그 이하에 위치한다. q90은 조건부 90% 분위값으로, 이론적으로 관측값의 약 90%가 그 이하에 위치한다. 따라서 q10 이하와 q90 초과에는 각각 약 10%의 관측값이 위치하며, q10과 q90 사이에는 조건부 분포의 중앙 약 80%가 포함된다.

이러한 분위 설정은 직경생장 예측 및 산림 생육범위 추정에 분위회귀식이 활용되어 온 선행연구를 고려해 반영한 것이다. 선행연구에서는 복수의 분위회귀식을 기반으로 직경생장 예측이 가능함을 제시하였으며, 99%와 같은 극단 분위수는 대규모 자료에서 이상치의 영향을 받을 수 있어 90% 분위수가 보다 안정적인 상위 경계 추정에 적합할 수 있다고 언급하였다(Bohora and Cao, 2014; Raptis et al., 2022). 또한 스페인 대규모 산림 인벤토리 자료를 활용한 연구에서는 생육범위의 경계를 설정하기 위해 수관폭의 90% 분위수와 추정수고의 10% 분위수를 기준으로 개방 생육목을 선별하고, 이들의 수관폭 관계를 분석하였다(Condés and Sterba, 2005). 이러한 선행연구를 바탕으로 본 연구에서는 극단 분위수에 비해 일부 관측값의 영향을 상대적으로 줄이면서 조건부 분포의 하위와 상위 위치를 함께 비교할 수 있도록 10%와 90% 분위수를 적용하였다.

수고-흉고직경 관계는 수종과 개체의 크기 범위에 따라 직선 또는 곡선 형태로 나타날 수 있으므로, 선형모형, 평균중심화 이차모형 및 cubic B-spline 모형을 후보모형으로 설정하였다. 선형모형은 실측수고가 서로 다른 개체들 사이에서 흉고직경의 조건부 분위값 차이가 일정한 비율로 나타난다고 가정하는 가장 단순한 모형이다. 회귀계수의 의미가 명확하고 수종 간 비교가 용이하지만, 수고 구간에 따라 흉고직경 분위값의 차이가 달라지는 관계를 충분히 표현하지 못할 수 있다.

이차모형은 실측수고와 그 제곱항을 사용하여 전체 수고 범위에 하나의 완만한 곡선을 적용하는 모형이다. 이 모형은 수고가 큰 개체에서 흉고직경 조건부 분위값의 차이가 점차 커지거나 작아지는 관계를 표현하기 위해 사용

하였다. 실측수고와 그 제곱항은 동일한 변수에서 생성되므로 구조적 공선성이 발생할 수 있다. 이를 완화하기 위해 각 개체목의 실측수고에서 해당 수종의 평균 실측수고를 뺀 평균중심화 변수와 그 제곱항을 사용하였다. 이때 1차항은 수종별 평균 수고 부근에서 수고가 서로 다른 개체들 사이에 나타나는 흉고직경 분위값의 차이를, 제곱항은 그 관계가 휘어지는 방향과 정도를 나타낸다.

Cubic B-spline 모형은 실측수고의 범위를 여러 구간으로 나누어 각 구간에 3차 다항식 형태의 곡선을 적용한 후, 구간 사이를 부드럽게 연결하는 방법이다. 이 모형은 낮은 수고 구간과 높은 수고 구간에 속한 개체들에서 흉고직경의 조건부 분위값 차이가 서로 다른 복잡한 비선형 관계를 검토하기 위해 사용하였다. 본 연구에서는 실측수고 범위에 균등하게 설정된 4개의 절점과 3차 B-spline을 적용하고, 표본점 단위 교차검증 예측성능을 선형모형 및 이차모형과 비교하였다. 즉, B-spline은 모든 수종의 기본모형이 아니라 선형 또는 이차모형보다 복잡한 함수형태가 실제로 필요한지를 확인하기 위한 비교 후보모형으로 사용하였다.

후보모형의 비교에는 표본점번호를 그룹으로 하는 5-fold 교차검증을 적용하였다. 교차검증의 fold 수는 모형 추정에 충분한 훈련자료를 확보하는 동시에 각 검증자료에도 일정 수 이상의 표본점이 포함되도록 하기 위해 5개로 설정하였다. 본 연구에서는 수종별 고유 표본점이 30개 이상인 경우를 분석대상으로 선정하였으므로, 5-fold 적용 시 각 반복에서 전체 표본점의 약 80%를 훈련자료로 활용하면서 나머지 약 20%를 독립적인 검증자료로 사용할 수 있다. 또한 최소 기준인 30개 표본점을 기준으로 할 경우 각 fold에 약 6개 표본점이 포함되어, 10-fold와 같이 fold 수를 증가시켰을 때 개별 검증 fold에 포함되는 표본점 수가 지나치게 감소하는 문제를 완화할 수 있다.

따라서 본 연구에서는 전체 표본점은 서로 겹치지 않는 5개 집단으로 나누고, 이 가운데 4개 집단으로 모형을 추정한 후 나머지 1개 집단으로 예측성능을 평가하였다. 검증 집단을 차례로 변경하여 동일한 과정을 5회 반복함으로써 모든 표본점이 한 번씩 검증자료로 사용되도록 하였다. 동일한 표본점에서 조사된 수목은 입지 및 임분환경 등 일부 생육조건을 공유할 수 있으므로, 같은 표본점의 관측값이 훈련자료와 검증자료에 동시에 포함될 경우 자료의 군집구조로 인해 예측성능이 과대평가될 가능성이 있다. 따라서 동일한 표본점에 속한 모든 개체목 관측값을 하나의 fold에 배정하여 훈련자료와 검증자료 간 표본점의 중복을 방지하였다. 이를 통해 모형 추정에 사용되지 않은 표본점의 개체목을 대상으로 각 후보모형이 목표 분위값을 얼마나 잘 추정하는지를 비교하였다.

후보모형의 예측오차는 분위회귀에 적합한 손실함수인 pinball loss를 이용하여 평가하였다. Pinball loss는 목표 분위에 따라 과대예측과 과소예측에 서로 다른 가중치를 부여한다. q10에서는 하위 분위값을 지나치게 높게 예측하는 경우에 상대적으로 큰 손실을 부여하고, q90에서는 상위 분위값을 지나치게 낮게 예측하는 경우에 상대적으로 큰 손실을 부여한다. 따라서 pinball loss가 작을수록 해당 모형이 검증자료에서 목표로 설정한 조건부 분위의 위치를 더 정확하게 추정한 것으로 해석하였다.

각 후보모형에서 q10과 q90의 교차검증 pinball loss를 각각 산출하고 두 값을 합산하여 하위 및 상위 분위의 예측성능을 함께 평가하였다. 이는 두 분위 가운데 한쪽만을 잘 추정하는 모형이 아니라, 흉고직경 조건부 분포의 하위와 상위 위치를 종합적으로 잘 추정하는 모형을 선정하기 위한 것이다.

최종모형은 합산 pinball loss에 1 표준오차 규칙을 적용하여 선정하였다. 먼저 평균 교차검증 손실이 가장 작은 모형을 확인하고, 해당 모형의 평균 손실에 1 표준오차를 더한 값을 허용 가능한 손실의 상한으로 설정하였다. 이후 이 범위에 포함되는 후보 중 선형모형, 이차모형, B-spline 모형의 순서로 가장 단순한 모형을 선택하였다. 이 절차는 복잡한 모형의 예측성능 개선이 교차검증 과정의 표본 변동을 고려한 이후에도 충분히 큰 경우에만 비선형 모형을 채택하고, 성능 차이가 크지 않은 경우에는 해석이 용이한 단순모형을 선택하기 위해 적용하였다.

최종모형의 성능은 교차검증 pinball loss, 관측비율, q10-q90 범위의 포괄률 및 분위회귀선의 교차 여부를 이용하여 평가하였다. 관측비율은 검증자료의 실제 흉고직경이 q10 또는 q90 예측값 이하에 위치하는 비율이며, 각각 목표 분위인 0.10과 0.90에 가까운지를 확인하였다. 포괄률은 실제 흉고직경이 q10과 q90 예측값 사이에 포함되는 비율로, 두 분위식이 의도한 조건부 중앙 80% 범위를 어느 정도 재현하는지를 확인하기 위해 사용하였다. 관측비율과 포괄률은 모형의 채택 또는 제외를 위한 절대적인 임계값이 아니라, 목표 분위와 분포 범위의 재현 정도를 확인하는 진단지표로 활용하였다.

회귀계수의 불확실성은 표본점 단위 군집 부트스트랩으로 평가하였다. 동일한 표본점의 수목은 생육환경을 공유하여 서로 유사할 수 있으므로, 개별 수목을 독립적으로 재표집하면 회귀계수의 불확실성이 실제보다 작게 평가될 가능성이 있다. 이에 개별 수목이 아닌 표본점을 복원추출하고, 선택된 표본점에 속한 모든 관측값을 함께 포함하여 최종모형을 다시 추정하였다. 이러한 재표집을 1,000회 반복하여 부트스트랩 표준오차와 백분위수 방식의 95% 신뢰구간을 산출하였다. 군집 부트스트랩은 표본점의 환경효과를 직접 제거하는 방법이 아니라, 동일 표본점의 관측값을 하나의 군집으로 유지하면서 표본점 구성에 따른 회귀계수의 불확실성을 평가하기 위해 적용하였다.

마지막으로 q10과 q90 분위회귀선이 조건부 분포의 순서를 일관되게 나타내는지를 확인하였다. q10은 하위 분위값이고 q90은 상위 분위값이므로, 동일한 수고에서 q10 예측값이 q90 예측값보다 크게 나타나면 두 분위식이 논리적인 순서를 충족하지 못한 것으로 볼 수 있다. 이에 수종별 관측 실측수고의 최솟값과 최댓값 사이에 500개의 평가 지점을 생성하고, 각 지점에서 두 분위회귀선의 교차 여부를 점검하였다. Pseudo $R^2$는 최종모형의 표본 내 상대적 적합도를 설명하기 위한 보조적 기술지표로만 산출하였으며, 특정 값을 모형의 채택 또는 제외 기준으로 사용하지 않았다.

## 2.4 분석 절차

본 연구에서는 분석자료 구축과 기초통계량 산정, 반복 관측구조 확인, 후보 분위회귀모형의 비교, 최종모형 추정과 평가 및 민감도 분석의 순서로 분석을 수행하였다. 먼저 2006–2020년 국가산림자원조사 원자료에서 수종명, 실측수고, 흉고직경, 집락번호, 표본점번호, 조사차기, 조사연도, 개체목 번호 및 개체목 구분코드를 추출하였다. 수종명은 표 1에 제시한 52개 후보 수종과 대조하여 일치하는 관측값만을 추출하였으며, 후보 수종 목록에 포함되지 않은 수종은 이후의 표본 수 기준 적용에 앞서 제외하였다. 실측수고 또는 흉고직경이 결측이거나 0 이하인 관측값도 분석에서 제외하였다. 이후 수종별 유효 관측값 수와 고유 표본점 수를 산정하고, 유효 관측값이 100개 이상이면서 고유 표본점이 30개 이상인 수종을 주 분석대상으로 선정하였다.

구축된 자료의 표본 규모와 기본적인 분포 특성을 파악하기 위해 전체 자료와 수종별 자료에서 실측수고와 흉고직경의 관측값 수, 평균, 표준편차, 중앙값, 최솟값 및 최댓값을 산출하였다. 또한 수종별 고유 표본점 수와 표본점당 관측값 수를 산정하여 관측값이 일부 표본점에 집중되어 있는지를 검토하였다. 최솟값과 최댓값은 입력 오류 또는 생물학적으로 해석하기 어려운 관측값의 존재 가능성을 확인하는 데 활용하였다. 이러한 기초통계량 검토를 거쳐 구축된 자료를 이후의 수종별 분위회귀분석에 사용하였다.

다음으로 제5차부터 제7차 국가산림자원조사 자료를 통합함에 따라 발생할 수 있는 관측값 간 종속성을 확인하였다. 표본점번호를 기준으로 둘 이상의 조사차기에 포함된 표본점을 반복 표본점으로 분류하였다. 동일 표본점에서 조사된 수목은 토양, 지형, 광환경 및 수분 조건 등의 생육환경을 공유할 수 있으므로, 교차검증에서는 같은 표본점의 관측값을 하나의 집단으로 배정하고 군집 부트스트랩에서는 표본점을 재표집 단위로 설정하였다. 또한 표본점번호와 개체목 번호를 결합하여 개체 후보 식별자를 구성하고, 동일한 식별자가 둘 이상의 조사차기에 나타난 경우 반복 개체 후보로 분류하였다. 반복 개체 후보는 동일한 개체가 통합자료에 여러 차례 포함되었을 가능성이 분석 결과에 미치는 영향을 확인하기 위한 민감도 분석에 활용하였다.

이후 흉고직경을 종속변수로, m 단위로 환산한 실측수고를 설명변수로 설정하여 수종별 q10 및 q90 분위회귀모형을 구축하였다. 각 분위에 대해 선형모형, 평균중심화 이차모형 및 cubic B-spline 모형을 후보로 설정하였다. 후보모형은 표본점번호를 그룹으로 하는 5-fold 교차검증으로 비교하였으며, 같은 표본점의 자료가 훈련자료와 검증자료에 동시에 포함되지 않도록 하였다. 각 fold에서 q10과 q90의 pinball loss를 산출하여 합산하고, 합산 손실의 평균과 표준오차를 이용한 1 표준오차 규칙을 적용하였다. 평균 합산 손실이 가장 작은 모형의 1 표준오차 범위에 포함되는 후보 중 가장 단순한 모형을 수종별 최종모형으로 선정하였다.

선정된 최종모형은 수종별 전체 분석자료에 다시 적합하여 q10 및 q90 분위회귀계수를 산출하였다. 회귀계수의 불확실성은 동일 표본점 내 관측값의 유사성을 고려할 수 있도록 표본점을 복원추출하는 군집 부트스트랩을 1,000회 수행하여 평가하였다. 부트스트랩 계수 분포를 이용하여 표준오차와 백분위수 방식의 95% 신뢰구간을 산출하였으며, 분석의 재현성을 위해 고정된 난수 시드를 적용하였다. 최종모형의 검증성능은 표본점 단위 교차검증에서 얻은 예측값을 이용하여 q10 및 q90의 관측비율과 두 분위 사이의 포괄률로 평가하였다. 관측비율은 각각 목표 분위인 0.10과 0.90에 가까운지를, 포괄률은 조건부 중앙 80% 범위를 적절하게 재현하는지를 확인하기 위해 사용하였다. 또한 수종별 관측 실측수고 범위에 500개의 평가 지점을 생성하여 q10 예측값이 q90 예측값보다 커지는 분위회귀선 교차 여부를 확인하였다.

마지막으로 최소 관측값 기준의 변화가 분석대상 수종에 미치는 영향과 반복 관측 가능성이 결과에 미치는 영향을 확인하기 위해 민감도 분석을 수행하였다. 최소 관측값 기준은 50개, 75개, 100개 및 150개로 변경하되, 고유 표본점 30개 이상의 기준은 동일하게 유지하였다. 고유 표본점 기준을 유지한 이유는 많은 개체목이 소수의 표본점에 집중되는 것을 방지하고, 표본점 단위 5-fold 교차검증과 군집 부트스트랩에 필요한 집단 수를 확보하는 동시에 최소 관측값 기준의 변화에 따른 영향만을 구분하여 확인하기 위함이다. 이에 따라 각 최소 관측값 기준에서 분석대상으로 포함되는 수종의 구성을 비교하였다. 이 분석은 최소 관측값 기준의 변화에 따른 분석대상 수종 구성의

민감도를 평가한 것으로, 개별 수종의 표본 수를 단계적으로 변화시키면서 q10 및 q90 회귀계수나 예측선의 안정성을 평가한 분석과는 구분하였다.

이와 별도로 반복 관측 가능성에 대해서는 반복 개체 후보별 최초 관측값 한 건만 남긴 자료로 모형을 다시 추정하고, 전체 통합자료에서 산출된 회귀계수와 방향 및 크기를 비교하였다. 아울러 제5차, 제6차 및 제7차 자료를 각각 분리하여 조사차기별 회귀계수를 추정하고, 이를 전체 통합자료의 회귀계수와 비교하였다.

# 3. 연구 결과

## 3.1 분석자료 구성 및 기초통계

표 1에 제시된 52개 1차 후보수종과 국가산림자원조사 원자료의 수종명을 대조하고, 실측수고와 흉고직경이 모두 유효한 관측값을 추출한 결과, 40개 수종의 64,382개 관측값이 확보되었다. 이 자료는 8,889개의 고유 표본점에서 조사되었으며, 제5차 22,878개, 제6차 20,752개 및 제7차 20,752개의 관측값으로 구성되었다.

수종별 유효 관측값 100개 이상과 고유 표본점 30개 이상의 기준을 적용한 결과, 25개 수종의 64,072개 관측값이 최종 분석자료에 포함되었다. 최종 분석자료는 8,856개 고유 표본점으로 구성되었다. 분석대상 수종 가운데 상수리나무의 관측값이 14,868개로 가장 많았고, 왕벚나무가 174개로 가장 적었다. 고유 표본점 수는 밤나무가 2,683개로 가장 많았으며, 백합나무가 31개로 가장 적었다.

최종 분석자료의 실측수고는 평균 11.7 ± 4.2m, 중앙값 11.4m였으며, 관측범위는 1.5-34.1m였다. 흉고직경은 평균 18.8 ± 9.0cm, 중앙값 17.0cm였으며, 관측범위는 6.0-104.0cm였다. 수종별 평균 실측수고는 함박꽃나무가 6.6m로 가장 낮고 은사시나무가 16.1m로 가장 높았다. 평균 흉고직경은 함박꽃나무가 9.7cm로 가장 작고 가래나무가 26.7cm로 가장 크게 나타났다. 수종별 최솟값과 최댓값을 검토한 결과, 양수가 아닌 값의 제외 이후 추가로 제거한 관측값은 없었다(표 2 참조).

표 2. 최종 분석대상 수종의 실측수고 및 흉고직경 기초통계

| 수종 | 관측값 | 표본점 | 수고 평균 ± SD(m) | 수고 중앙값(m) | 수고 범위(m) | DBH 평균 ± SD(cm) | DBH 중앙값(cm) |
|---|---|---|---|---|---|---|---|
| 전체 | 64,072 | 8,856 | 11.7 ± 4.2 | 11.4 | 1.5-34.1 | 18.8 ± 9.0 | 17.0 |
| 가래나무 | 1,493 | 341 | 15.1 ± 4.8 | 14.7 | 3.6-30.2 | 26.7 ± 11.4 | 25.0 |
| 가죽나무 | 224 | 82 | 10.9 ± 3.4 | 11.0 | 3.2-21.3 | 14.3 ± 6.3 | 13.0 |
| 갈참나무 | 8,897 | 2,226 | 11.6 ± 3.6 | 11.4 | 2.0-26.7 | 17.3 ± 7.8 | 16.0 |
| 감나무 | 289 | 103 | 7.6 ± 2.7 | 7.3 | 2.4-15.9 | 13.7 ± 5.9 | 13.0 |
| 고로쇠나무 | 4,030 | 1,123 | 11.5 ± 3.3 | 11.4 | 3.0-25.2 | 18.6 ± 9.1 | 17.0 |
| 느티나무 | 1,337 | 335 | 11.6 ± 3.5 | 11.4 | 3.2-24.1 | 19.1 ± 9.5 | 17.0 |
| 단풍나무 | 254 | 76 | 8.3 ± 3.1 | 8.0 | 2.1-17.9 | 13.4 ± 6.2 | 11.0 |
| 당단풍나무 | 2,311 | 796 | 8.1 ± 2.0 | 8.0 | 2.4-21.0 | 11.7 ± 4.3 | 11.0 |
| 마가목 | 222 | 119 | 8.9 ± 2.8 | 8.9 | 2.0-15.3 | 15.5 ± 7.2 | 14.0 |
| 밤나무 | 13,872 | 2,683 | 10.5 ± 3.7 | 10.3 | 1.8-26.9 | 18.9 ± 8.9 | 17.0 |
| 백합나무 | 239 | 31 | 10.1 ± 3.9 | 9.2 | 3.5-20.7 | 12.8 ± 6.3 | 10.0 |
| 버드나무 | 1,120 | 272 | 10.6 ± 3.5 | 10.5 | 2.6-21.4 | 19.1 ± 9.5 | 17.0 |
| 뽕나무 | 226 | 67 | 7.6 ± 2.8 | 7.2 | 2.5-18.0 | 12.0 ± 6.6 | 10.0 |
| 산딸나무 | 474 | 195 | 8.2 ± 2.3 | 8.0 | 3.2-16.7 | 12.9 ± 5.3 | 11.0 |
| 상수리나무 | 14,868 | 2,606 | 14.2 ± 4.2 | 14.1 | 2.2-31.6 | 21.3 ± 8.6 | 20.0 |
| 왕벚나무 | 174 | 44 | 10.7 ± 3.3 | 10.4 | 4.1-20.2 | 22.0 ± 9.6 | 21.5 |
| 은사시나무 | 809 | 169 | 16.1 ± 5.8 | 15.0 | 4.1-34.1 | 20.1 ± 10.2 | 18.0 |
| 자귀나무 | 1,194 | 589 | 9.2 ± 2.8 | 9.1 | 2.5-19.4 | 14.0 ± 5.4 | 13.0 |
| 자작나무 | 733 | 90 | 10.3 ± 3.4 | 9.9 | 4.0-24.6 | 11.9 ± 4.7 | 11.0 |

표 2. 계속

| 수종 | 관측값 | 표본점 | 수고 평균 ± SD (m) | 수고 중앙값 (m) | 수고 범위 (m) | DBH 평균 ± SD (cm) | DBH 중앙값 (cm) |
|---|---|---|---|---|---|---|---|
| 쪽동백나무 | 1,796 | 683 | 8.4 ± 2.0 | 8.3 | 2.4–19.0 | 11.6 ± 3.9 | 11.0 |
| 층층나무 | 4,373 | 1,147 | 12.3 ± 3.3 | 12.4 | 2.4–27.8 | 20.9 ± 8.8 | 20.0 |
| 팽나무 | 1,873 | 559 | 10.4 ± 3.6 | 10.1 | 1.5–25.5 | 18.6 ± 9.0 | 17.0 |
| 피나무 | 2,530 | 606 | 12.6 ± 3.6 | 12.4 | 3.5–28.7 | 20.4 ± 10.5 | 18.0 |
| 함박꽃나무 | 418 | 166 | 6.6 ± 1.8 | 6.6 | 3.0–14.7 | 9.7 ± 2.8 | 9.0 |
| 황벽나무 | 316 | 129 | 12.3 ± 3.6 | 12.1 | 3.5–22.2 | 20.5 ± 7.6 | 19.0 |

### 3.2 반복 표본점 및 반복 개체 후보의 분포

최종 분석자료에 포함된 8,856개 고유 표본점의 조사차기별 반복 여부를 확인한 결과, 2개 이상의 조사차기에 포함된 반복 표본점은 6,529개로 전체 표본점의 73.7%를 차지하였다. 이 가운데 1,664개 표본점은 2개 조사차기에, 4,865개 표본점은 제5차부터 제7차까지 3개 조사차기에 모두 포함되었다. 한 개 조사차기에만 포함된 표본점은 2,327개로 나타났다.

분석대상 수종 전체에서 표본점번호와 개체목 번호를 결합하여 구성한 개체 후보 식별자는 총 56,651개였다. 이 가운데 2개 이상의 조사차기에 동일한 식별자가 나타난 반복 개체 후보는 6,464개로 전체 개체 후보의 11.4%를 차지하였다. 반복 개체 후보 중 5,646개는 2개 조사차기에, 818개는 3개 조사차기에 나타났다. 다만 이 결과는 공개 원자료에서 동일한 표본점번호와 개체목 번호가 반복된 경우를 집계한 것으로, 조사차기 사이에 동일한 개체목임을 확정한 결과가 아니라 반복 가능성이 있는 개체 후보의 분포를 나타낸다(표 3 참조).

### 3.3 후보모형 비교 및 최종모형 선정

최종 분석대상 25개 수종에 대해 선형모형, 평균중심화 이차모형 및 cubic B-spline 모형의 표본점 단위 5-fold 교차검증을 수행한 결과, 모든 수종과 후보모형에서 실패한 fold 없이 q10 및 q90 합산 pinball loss가 산출되었다.

평균 합산 pinball loss가 가장 작은 모형만을 기준으로 비교하면 선형모형이 10개 수종, 이차모형이 9개 수종, B-spline 모형이 6개 수종에서 가장 낮은 손실을 나타냈다. 그러나 최소 손실 모형의 1 표준오차 범위에 포함되는 후보 중 가장 단순한 모형을 선택한 결과, 최종적으로 23개 수종에서 선형모형이 선정되었고 가죽나무와 갈참나무에서는 이차모형이 선정되었다. B-spline 모형이 최종모형으로 선정된 수종은 없었다.

가죽나무의 평균 합산 pinball loss는 이차모형에서 1.3456으로 가장 낮았으며, 1 표준오차 허용 상한은 1.4311이었다. 선형모형의 평균 손실은 1.4401로 허용 상한보다 높게 나타나 이차모형이 최종모형으로 선정되었다. 갈참나무에서도 이차모형의 평균 손실이 1.7328로 가장 낮았으며, 허용 상한은 1.7482였다. 선형모형의 평균 손실은 1.7494로 허용 상한을 초과하여 이차모형이 선정되었다.

느티나무, 마가목, 밤나무, 상수리나무, 은사시나무 및 층층나무에서는 B-spline 모형의 평균 손실이 가장 작게 나타났으나, 선형모형의 평균 손실도 최소 손실 모형의 1 표준오차 범위에 포함되었다. 이에 따라 해당 수종에서는 예측성능의 차이가 모형의 복잡성 증가를 뒷받침할 정도로 크지 않은 것으로 판단하여 선형모형을 선정하였다. 가래나무, 고로쇠나무, 당단풍나무, 백합나무, 쪽동백나무, 팽나무 및 함박꽃나무에서도 이차모형의 평균 손실이 가장 작았으나 선형모형이 1 표준오차 범위에 포함되어 최종적으로 선형모형이 선정되었다(표 4 참조).

표 3. 조사차기별 표본점 및 개체 후보의 반복 분포

| 관측된 조사차기 수 | 표본점 수 | 표본점 비율(%) | 개체 후보 수 |
|---|---|---|---|
| 1개 조사차기 | 2,327 | 26.3 | 50,187 |
| 2개 조사차기 | 1,664 | 18.8 | 5,646 |
| 3개 조사차기 | 4,865 | 54.9 | 818 |
| 전체 | 8,856 | 100 | 56,651 |

표 4. 후보모형의 평균 합산 핀볼 로스와 최종모형 선정 결과

| 수종 | 선형 | 이차 | B-spline | 최소손실 | 1 SE 상한 | 최종 |
|---|---|---|---|---|---|---|
| 가래나무 | 2.8392 | 2.8127 | 2.8235 | 이차 | 2.9285 | 선형 |
| 가죽나무 | 1.4391 | 1.3456 | 1.3660 | 이차 | 1.4311 | 이차 |
| 갈참나무 | 1.7494 | 1.7328 | 1.7359 | 이차 | 1.7482 | 이차 |
| 감나무 | 1.5184 | 1.5654 | 1.7423 | 선형 | 1.7037 | 선형 |
| 고로쇠나무 | 2.3337 | 2.3261 | 2.3340 | 이차 | 2.4029 | 선형 |
| 느티나무 | 2.2231 | 2.2281 | 2.2125 | B-spline | 2.3929 | 선형 |
| 단풍나무 | 1.4653 | 1.4736 | 1.8315 | 선형 | 1.5970 | 선형 |
| 당단풍나무 | 1.1772 | 1.1693 | 1.1925 | 이차 | 1.2023 | 선형 |
| 마가목 | 2.0240 | 2.0254 | 1.9840 | B-spline | 2.2255 | 선형 |
| 밤나무 | 2.3353 | 2.3307 | 2.3273 | B-spline | 2.3730 | 선형 |
| 백합나무 | 0.9410 | 0.9258 | 1.0668 | 이차 | 0.9772 | 선형 |
| 버드나무 | 2.4680 | 2.4743 | 2.4888 | 선형 | 2.7096 | 선형 |
| 뽕나무 | 1.5541 | 1.6210 | 1.9485 | 선형 | 1.9571 | 선형 |
| 산딸나무 | 1.4853 | 1.5738 | 1.6748 | 선형 | 1.5482 | 선형 |
| 상수리나무 | 2.0765 | 2.0727 | 2.0653 | B-spline | 2.0944 | 선형 |
| 왕벚나무 | 2.7751 | 2.9100 | 3.7574 | 선형 | 3.1279 | 선형 |
| 은사시나무 | 2.1794 | 2.2106 | 2.1570 | B-spline | 2.2613 | 선형 |
| 자귀나무 | 1.5195 | 1.5225 | 1.5267 | 선형 | 1.5929 | 선형 |
| 자작나무 | 1.0460 | 1.0650 | 1.1007 | 선형 | 1.1253 | 선형 |
| 쪽동백나무 | 1.0606 | 1.0501 | 1.0504 | 이차 | 1.0997 | 선형 |
| 층층나무 | 2.2432 | 2.2408 | 2.2373 | B-spline | 2.2816 | 선형 |
| 팽나무 | 2.1847 | 2.1707 | 2.1913 | 이차 | 2.2704 | 선형 |
| 피나무 | 2.6237 | 2.6282 | 2.6254 | 선형 | 2.6991 | 선형 |
| 함박꽃나무 | 0.9025 | 0.8841 | 0.9395 | 이차 | 0.9677 | 선형 |
| 황벽나무 | 1.9194 | 1.9377 | 2.0206 | 선형 | 2.0647 | 선형 |

## 3.4 수종별 최종 분위회귀식 및 계수 불확실성

1 표준오차 규칙에 따라 선정된 수종별 최종모형을 각 수종의 전체 분석자료에 다시 적합하여 q10 및 q90 분위 회귀계수를 추정하였다(표 5 참조). H는 현장에서 측정한 수고(m), DBH는 흉고직경(cm)을 의미한다.

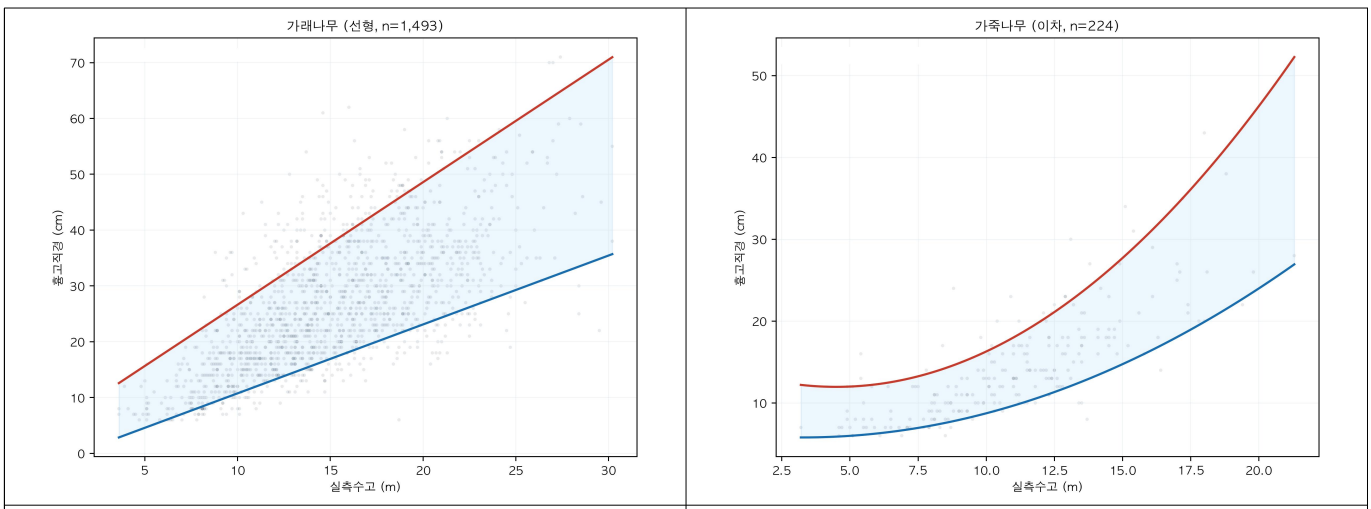
선형모형은 $DBH = \beta 0 + \beta 1 H$로, 이차모형은 $DBH = \beta 0 + \beta 1 (H - \overline{H}) + \beta 2 (H - \overline{H})^2$로 표현하였다. 표 5에 제시한 수고 범위는 각 수종의 최종 분석자료에서 관측된 실측수고의 최솟값과 최댓값이며, 이를 벗어난 수고값에 대한 예측은 외삽에 해당하므로 회귀식은 해당 범위 내에서만 적용하였다. 수종별 최종 분위회귀선은 그림 1에 제시하였다. 회귀계수의 불확실성은 표본점을 재표집 단위로 하는 군집 부트스트랩을 1,000회 수행하여 평가하였으며, 부트스트랩 표준오차와 백분위수 방식의 95% 신뢰구간을 산출하였다(표 6 참조).

최종모형의 1차항 추정치는 모든 수종의 q10과 q90에서 양수로 나타났다. 이는 관측된 실측수고 범위에서 수고가 큰 개체일수록 흉고직경의 조건부 하위 및 상위 분위값도 전반적으로 크게 나타났음을 의미한다. 또한 모든 수종에서 q90의 1차항 추정치가 q10보다 컸다. 따라서 개체 간 실측수고 차이에 대응하는 흉고직경 분위값의 차이는 조건부 분포의 하위보다 상위에서 상대적으로 크게 나타났다. 다만 함박꽃나무는 q10과 q90 모두 1차항의 군집 부트스트랩 신뢰구간 하한이 0에 근접하여 다른 수종보다 수고-흉고직경 관계의 불확실성이 컸다.

이차모형이 선정된 가죽나무와 갈참나무에서는 q10과 q90의 이차항이 모두 양수로 추정되었으며, 군집 부트스트랩 95% 신뢰구간도 0보다 큰 범위로 나타났다. 가죽나무의 이차항은 q10에서 0.065, q90에서 0.143이었고, 갈

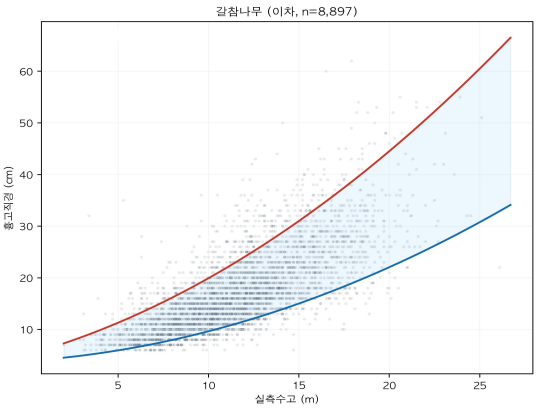
표 5. 수종별 q10 및 q90 분위회귀식과 적용 가능한 실측수고 범위

| 수종 | 모형 | 수고 범위(m) | q10 회귀식 | q90 회귀식 |
|---|---|---|---|---|
| 가래나무 | 선형 | 3.6–30.2 | q10: DBH = −1.593 + 1.235H | q90: DBH = 4.683 + 2.195H |
| 가죽나무 | 이차 | 3.2–21.3 | q10: DBH = 9.529 + 0.987(H − 10.856) + 0.065(H − 10.856)$^2$ | q90: DBH = 17.736 + 1.816(H − 10.856) + 0.143(H − 10.856)$^2$ |
| 갈참나무 | 이차 | 2.0–26.7 | q10: DBH = 11.217 + 1.017(H − 11.637) + 0.033(H − 11.637)$^2$ | q90: DBH = 23.279 + 2.131(H − 11.637) + 0.049(H − 11.637)$^2$ |
| 감나무 | 선형 | 2.4–15.9 | q10: DBH = 2.455 + 0.909H | q90: DBH = 4.000 + 2.000H |
| 고로쇠나무 | 선형 | 3.0–25.2 | q10: DBH = −0.529 + 1.059H | q90: DBH = −0.619 + 2.381H |
| 느티나무 | 선형 | 3.2–24.1 | q10: DBH = −0.043 + 1.087H | q90: DBH = −0.767 + 2.427H |
| 단풍나무 | 선형 | 2.1–17.9 | q10: DBH = 2.162 + 0.811H | q90: DBH = 2.000 + 2.000H |
| 당단풍나무 | 선형 | 2.4–21.0 | q10: DBH = 3.235 + 0.588H | q90: DBH = 2.522 + 1.739H |
| 마가목 | 선형 | 2.0–15.3 | q10: DBH = 3.154 + 0.769H | q90: DBH = 4.167 + 2.083H |
| 밤나무 | 선형 | 1.8–26.9 | q10: DBH = 0.655 + 1.034H | q90: DBH = 9.045 + 1.791H |
| 백합나무 | 선형 | 3.5–20.7 | q10: DBH = −1.015 + 1.029H | q90: DBH = 0.078 + 1.638H |
| 버드나무 | 선형 | 2.6–21.4 | q10: DBH = 1.279 + 0.984H | q90: DBH = 3.829 + 2.190H |
| 뽕나무 | 선형 | 2.5–18.0 | q10: DBH = 3.600 + 0.500H | q90: DBH = 0.243 + 2.214H |
| 산딸나무 | 선형 | 3.2–16.7 | q10: DBH = 4.318 + 0.455H | q90: DBH = 3.985 + 1.846H |
| 상수리나무 | 선형 | 2.2–31.6 | q10: DBH = −0.444 + 1.056H | q90: DBH = 3.412 + 1.824H |
| 왕벚나무 | 선형 | 4.1–20.2 | q10: DBH = −1.714 + 1.429H | q90: DBH = 1.985 + 2.687H |
| 은사시나무 | 선형 | 4.1–34.1 | q10: DBH = −2.644 + 1.017H | q90: DBH = −0.826 + 1.739H |
| 자귀나무 | 선형 | 2.5–19.4 | q10: DBH = 3.312 + 0.625H | q90: DBH = 6.493 + 1.467H |
| 자작나무 | 선형 | 4.0–24.6 | q10: DBH = 1.256 + 0.698H | q90: DBH = 1.552 + 1.379H |
| 쪽동백나무 | 선형 | 2.4–19.0 | q10: DBH = 2.143 + 0.714H | q90: DBH = 4.305 + 1.356H |
| 층층나무 | 선형 | 2.4–27.8 | q10: DBH = −2.000 + 1.250H | q90: DBH = 2.824 + 2.157H |
| 팽나무 | 선형 | 1.5–25.5 | q10: DBH = −0.230 + 1.148H | q90: DBH = 3.381 + 2.268H |
| 피나무 | 선형 | 3.5–28.7 | q10: DBH = −0.717 + 1.087H | q90: DBH = −2.981 + 2.547H |
| 함박꽃나무 | 선형 | 3.0–14.7 | q10: DBH = 4.774 + 0.323H | q90: DBH = 9.514 + 0.571H |
| 황벽나무 | 선형 | 3.5–22.2 | q10: DBH = −2.400 + 1.333H | q90: DBH = 8.850 + 1.500H |

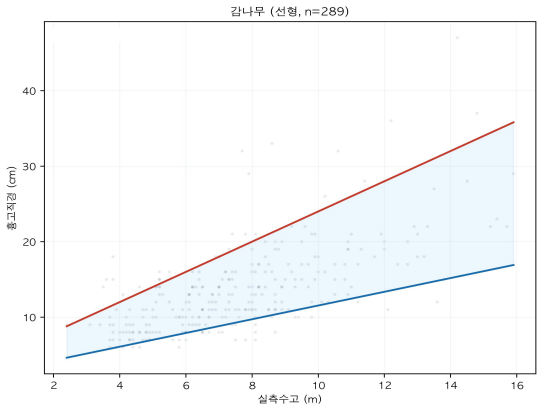



a: 가래나무 분위회귀선(선형모형) | b: 가죽나무 분위회귀선(이차모형)

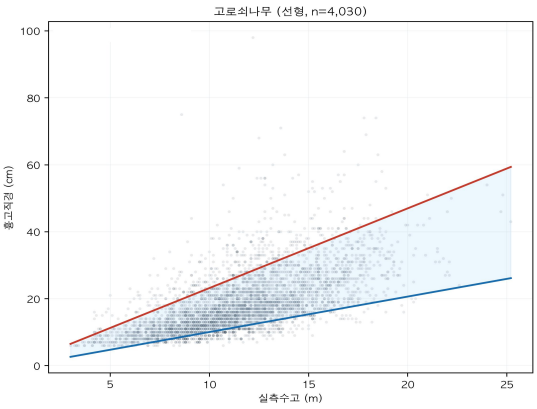
그림 1. 최종 수종별 분위회귀선
범례: q10 q90



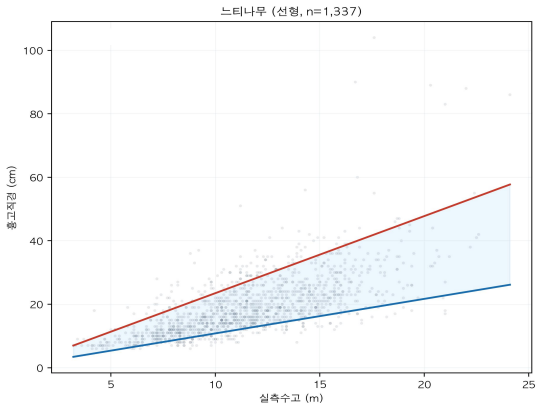

c: 갈참나무 분위회귀선(이차모형)



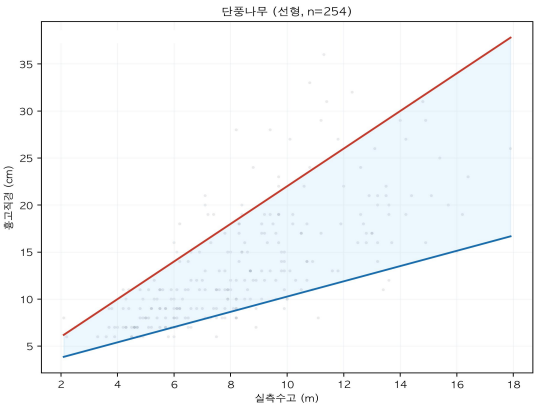

d: 감나무 분위회귀선(선형모형)



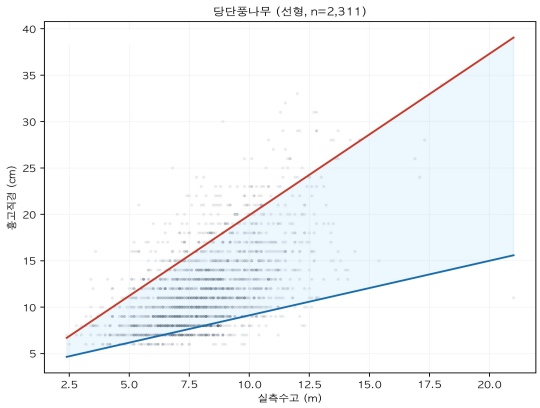

e: 고로쇠나무 분위회귀선(선형모형)



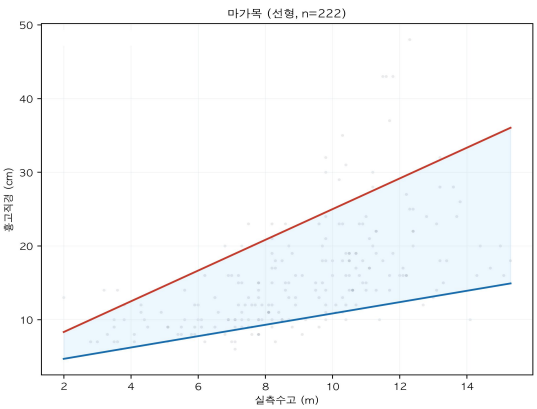

f: 느티나무 분위회귀선(선형모형)



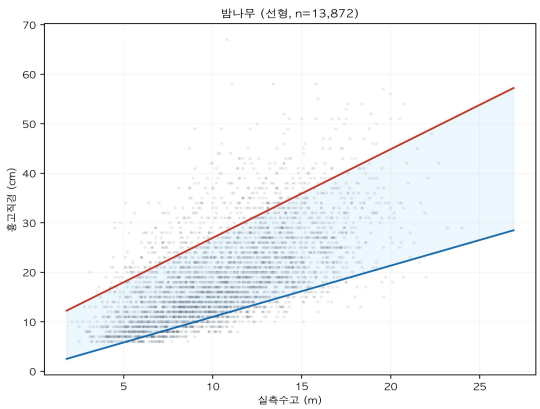

g: 단풍나무 분위회귀선(선형모형)




h: 당단풍나무 분위회귀선(선형모형)

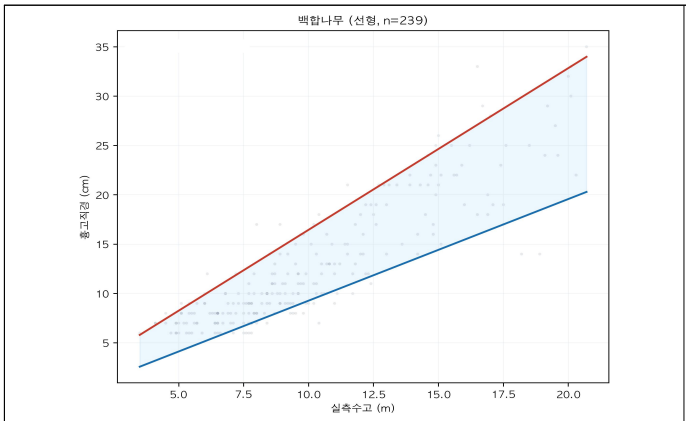



i: 마가목 분위회귀선(선형모형)

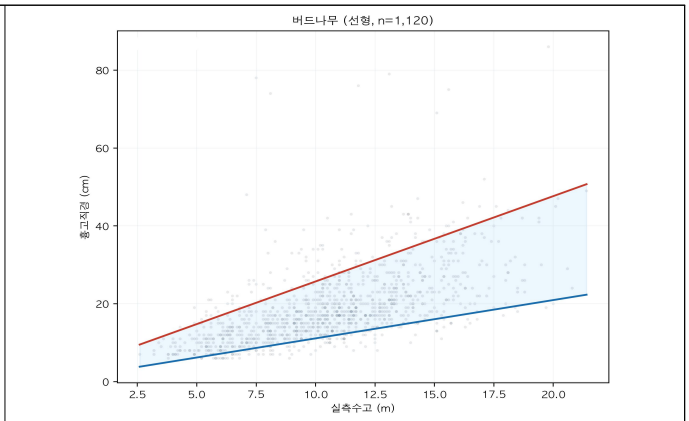



j: 밤나무 분위회귀선(선형모형)

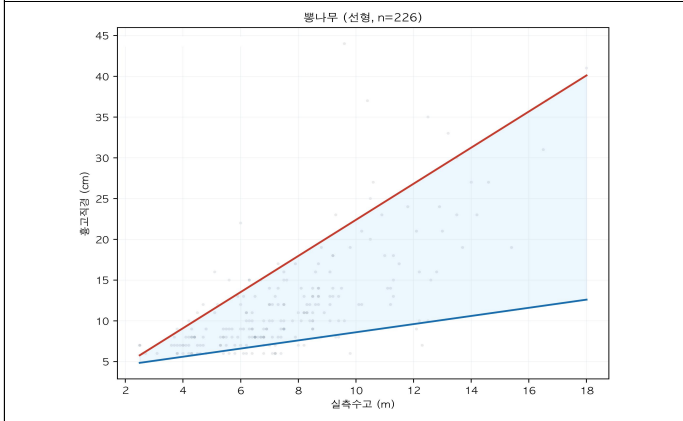
그림 1. 계속



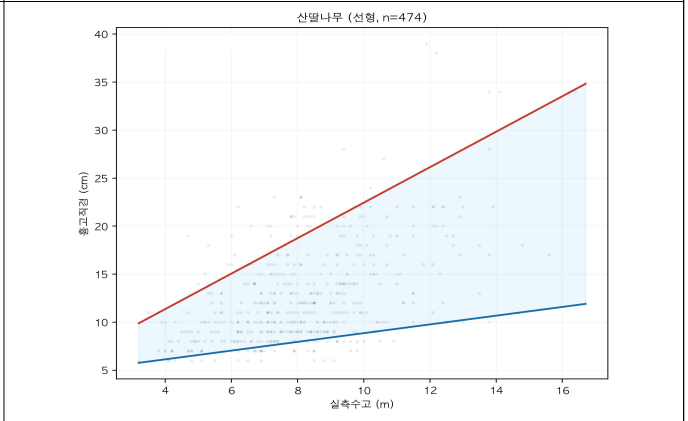

k: 백합나무 분위회귀선(선형모형)



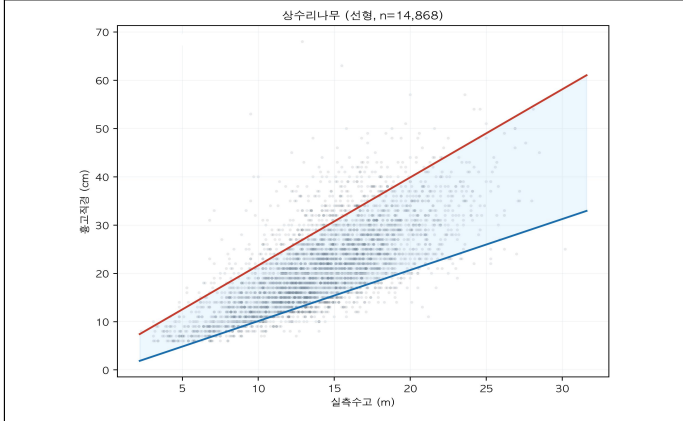

l: 버드나무 분위회귀선(선형모형)



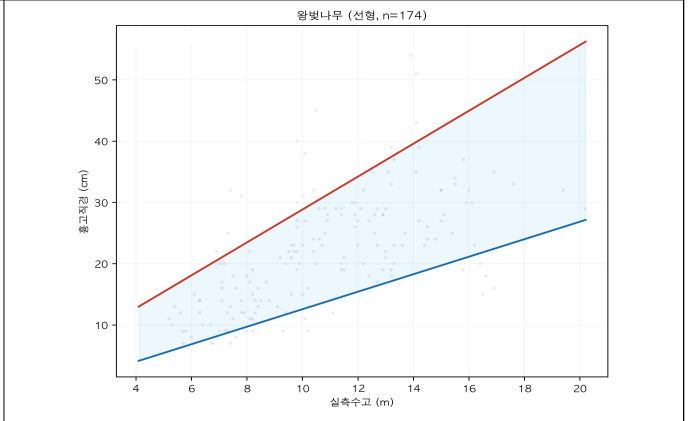

m: 뽕나무 분위회귀선(선형모형)



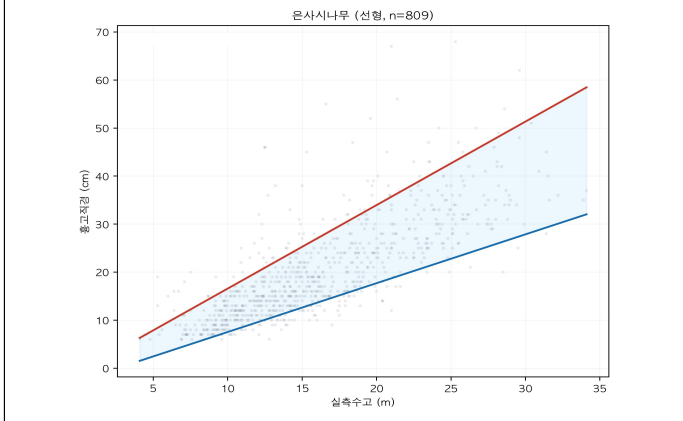

n: 산딸나무 분위회귀선(선형모형)



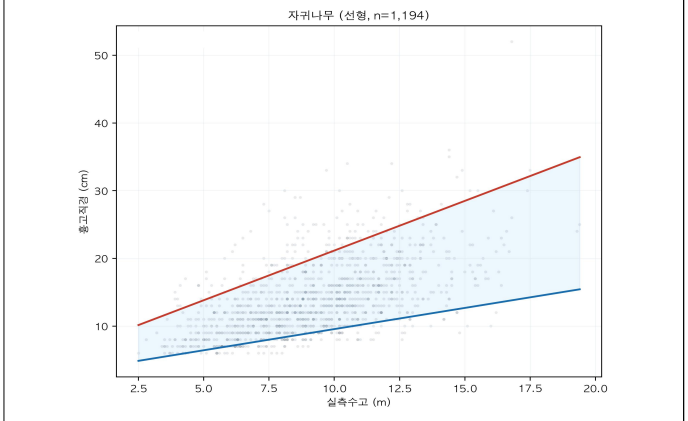

o: 상수리나무 분위회귀선(선형모형)




p: 왕벚나무 분위회귀선(선형모형)

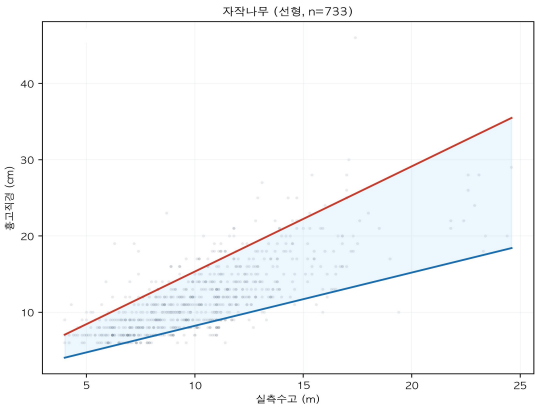



q: 은사시나무 분위회귀선(선형모형)

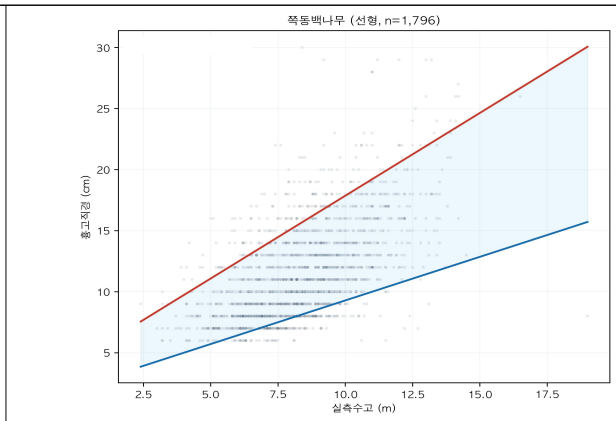



r: 자귀나무 분위회귀선(선형모형)

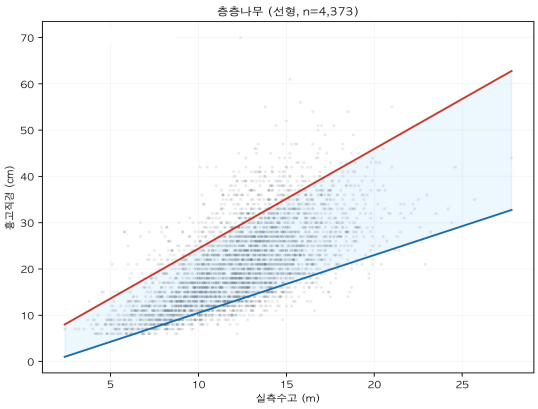
그림 1. 계속



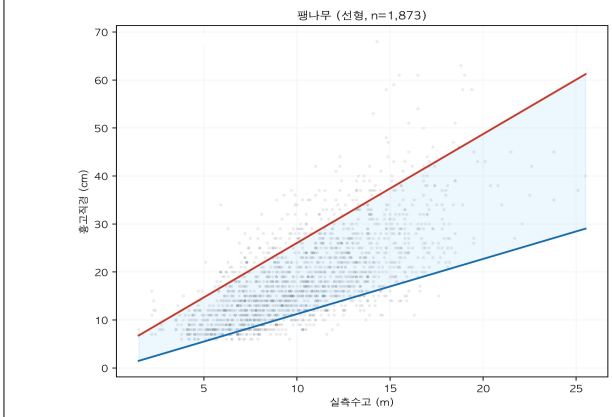

s: 자작나무 분위회귀선(선형모형)



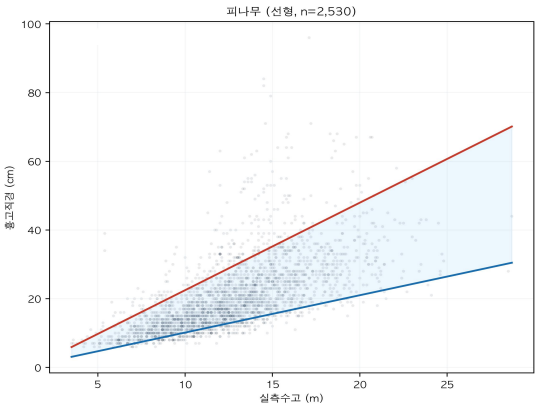

t: 쪽동백나무 분위회귀선(선형모형)



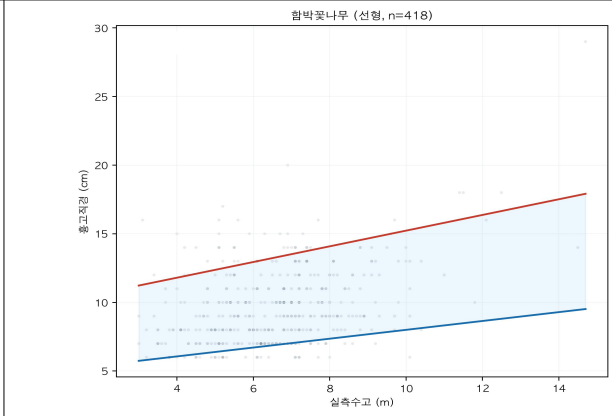

u: 층층나무 분위회귀선(선형모형)



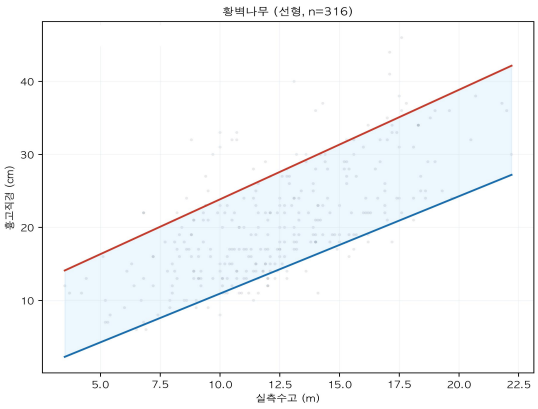

v: 팽나무 분위회귀선(선형모형)




w: 피나무 분위회귀선(선형모형)




x: 함박꽃나무 분위회귀선(선형모형)




y: 황벽나무 분위회귀선(선형모형)

그림 1. 계속

표 6. 최종모형 회귀계수와 표본점 단위 군집 부트스트랩 95% 신뢰구간

| 수종 | 분위 | 모형 | Pseudo $R^2$ | 절편 [95% 신뢰구간] | 1차계수 [95% 신뢰구간] | 2차계수 [95% 신뢰구간] |
|---|---|---|---|---|---|---|
| 가래나무 | q10 | 선형 | 0.305 | −1.593 [−2.996, 0.547] | 1.235 [1.087, 1.379] | - |
| | q90 | 선형 | 0.256 | 4.683 [0.715, 12.845] | 2.195 [1.702, 2.463] | - |
| 가죽나무 | q10 | 이차 | 0.358 | 9.529 [9.235, 10.606] | 0.987 [0.907, 1.191] | 0.065 [0.031, 0.086] |
| | q90 | 이차 | 0.435 | 17.736 [16.173, 19.403] | 1.816 [1.238, 2.277] | 0.143 [0.002, 0.244] |
| 갈참나무 | q10 | 이차 | 0.283 | 11.217 [11.086, 11.431] | 1.017 [0.974, 1.068] | 0.033 [0.018, 0.046] |
| | q90 | 이차 | 0.376 | 23.279 [22.736, 23.741] | 2.131 [2.001, 2.219] | 0.049 [0.030, 0.065] |
| 감나무 | q10 | 선형 | 0.186 | 2.455 [1.158, 4.236] | 0.909 [0.588, 1.111] | - |
| | q90 | 선형 | 0.315 | 4.000 [0.418, 9.410] | 2.000 [1.142, 2.500] | - |
| 고로쇠나무 | q10 | 선형 | 0.214 | −0.529 [−1.308, 0.336] | 1.059 [0.972, 1.143] | - |
| | q90 | 선형 | 0.244 | −0.619 [−2.390, 2.305] | 2.381 [2.121, 2.568] | - |
| 느티나무 | q10 | 선형 | 0.267 | −0.043 [−1.318, 1.288] | 1.087 [0.930, 1.212] | - |
| | q90 | 선형 | 0.285 | −0.767 [−4.098, 2.590] | 2.427 [2.114, 2.727] | - |
| 단풍나무 | q10 | 선형 | 0.239 | 2.162 [0.333, 3.067] | 0.811 [0.659, 1.064] | - |
| | q90 | 선형 | 0.353 | 2.000 [−2.914, 4.579] | 2.000 [1.600, 2.882] | - |
| 당단풍나무 | q10 | 선형 | 0.103 | 3.235 [2.652, 4.000] | 0.588 [0.484, 0.667] | - |
| | q90 | 선형 | 0.231 | 2.522 [1.273, 4.164] | 1.739 [1.500, 1.899] | - |
| 마가목 | q10 | 선형 | 0.175 | 3.154 [−0.154, 4.684] | 0.769 [0.556, 1.111] | - |
| | q90 | 선형 | 0.193 | 4.167 [−2.100, 10.083] | 2.083 [1.286, 3.294] | - |
| 밤나무 | q10 | 선형 | 0.238 | 0.655 [0.245, 1.000] | 1.034 [1.000, 1.081] | - |
| | q90 | 선형 | 0.199 | 9.045 [7.193, 10.913] | 1.791 [1.636, 1.954] | - |
| 백합나무 | q10 | 선형 | 0.334 | −1.015 [−2.207, 0.824] | 1.029 [0.811, 1.143] | - |
| | q90 | 선형 | 0.644 | 0.078 [−1.746, 1.027] | 1.638 [1.457, 1.789] | - |
| 버드나무 | q10 | 선형 | 0.227 | 1.279 [−0.577, 2.565] | 0.984 [0.857, 1.184] | - |
| | q90 | 선형 | 0.244 | 3.829 [0.830, 7.190] | 2.190 [1.940, 2.500] | - |
| 뽕나무 | q10 | 선형 | 0.107 | 3.600 [−1.176, 5.049] | 0.500 [0.250, 1.250] | - |
| | q90 | 선형 | 0.465 | 0.243 [−2.097, 3.729] | 2.214 [1.568, 2.624] | - |
| 산딸나무 | q10 | 선형 | 0.085 | 4.318 [2.674, 5.256] | 0.455 [0.357, 0.698] | - |
| | q90 | 선형 | 0.220 | 3.985 [0.399, 8.200] | 1.846 [1.250, 2.322] | - |
| 상수리나무 | q10 | 선형 | 0.332 | −0.444 [−0.890, −0.051] | 1.056 [1.024, 1.094] | - |
| | q90 | 선형 | 0.267 | 3.412 [2.504, 4.650] | 1.824 [1.739, 1.894] | - |
| 왕벚나무 | q10 | 선형 | 0.228 | −1.714 [−8.680, 3.200] | 1.429 [0.952, 2.206] | - |
| | q90 | 선형 | 0.222 | 1.985 [−5.384, 18.097] | 2.687 [1.288, 3.333] | - |
| 은사시나무 | q10 | 선형 | 0.353 | −2.644 [−5.293, −0.577] | 1.017 [0.857, 1.205] | - |
| | q90 | 선형 | 0.371 | −0.826 [−3.645, 6.359] | 1.739 [1.357, 1.977] | - |
| 자귀나무 | q10 | 선형 | 0.160 | 3.312 [1.999, 3.833] | 0.625 [0.556, 0.800] | - |
| | q90 | 선형 | 0.180 | 6.493 [4.000, 9.543] | 1.467 [1.111, 1.705] | - |
| 자작나무 | q10 | 선형 | 0.216 | 1.256 [0.462, 2.597] | 0.698 [0.556, 0.769] | - |
| | q90 | 선형 | 0.384 | 1.552 [0.250, 4.643] | 1.379 [1.064, 1.494] | - |
| 쪽동백나무 | q10 | 선형 | 0.130 | 2.143 [1.083, 3.371] | 0.714 [0.571, 0.834] | - |
| | q90 | 선형 | 0.218 | 4.305 [2.692, 6.111] | 1.356 [1.154, 1.538] | - |
| 층층나무 | q10 | 선형 | 0.272 | −2.000 [−2.981, −1.441] | 1.250 [1.194, 1.346] | - |
| | q90 | 선형 | 0.227 | 2.824 [0.059, 5.472] | 2.157 [1.944, 2.400] | - |
| 팽나무 | q10 | 선형 | 0.229 | −0.230 [−1.250, 1.327] | 1.148 [0.952, 1.250] | - |
| | q90 | 선형 | 0.322 | 3.381 [1.400, 6.267] | 2.268 [2.000, 2.522] | - |
| 피나무 | q10 | 선형 | 0.262 | −0.717 [−2.334, 0.275] | 1.087 [0.988, 1.220] | - |
| | q90 | 선형 | 0.254 | −2.981 [−5.747, −0.766] | 2.547 [2.326, 2.857] | - |
| 함박꽃나무 | q10 | 선형 | 0.028 | 4.774 [3.953, 7.000] | 0.323 [0.000, 0.476] | - |
| | q90 | 선형 | 0.056 | 9.514 [6.073, 14.000] | 0.571 [0.000, 1.020] | - |
| 황벽나무 | q10 | 선형 | 0.305 | −2.400 [−4.185, 0.308] | 1.333 [1.111, 1.481] | - |
| | q90 | 선형 | 0.257 | 8.850 [4.870, 14.889] | 1.500 [1.071, 1.809] | - |

참나무는 각각 0.033과 0.049였다. 이는 두 수종에서 수고가 큰 개체군으로 갈수록 흉고직경 분위값의 개체 간 차이가 일정하게 유지되기보다 커지는 곡선 관계가 나타났음을 의미한다. 이차모형의 계수는 수종별 평균 실측수고를 기준으로 중심화하여 산출하였으며, 중심수고는 가죽나무 10.856m, 갈참나무 11.637m였다.

### 3.5 최종모형의 교차검증 성능 및 분위선 교차

수종별 최종모형의 검증성능은 표본점 단위 5-fold 교차검증에서 산출된 예측값을 이용하여 평가하였다(표 7 참조). 모든 수종에서 5개의 교차검증 fold가 정상적으로 분석되어 모형 추정에 실패한 fold는 나타나지 않았다.

q10의 관측비율은 0.095–0.132의 범위로 나타났으며, 수종 평균은 0.105였다. q90의 관측비율은 0.885–0.906의 범위였고, 전체 평균은 0.896으로 나타났다. 따라서 검증자료에서 q10과 q90의 관측비율은 각각 목표 분위인 0.10과 0.90에 전반적으로 근접하였다.

q10과 q90 예측값 사이에 실제 흉고직경이 포함된 비율인 포괄률은 0.759–0.806의 범위로 나타났으며, 수종 평균은 0.791이었다. 이는 이론적인 중앙 80% 포괄률인 0.800에 근접한 결과이다. 자작나무의 포괄률이 0.806으로 가장 높았으며, 왕벚나무는 0.759로 가장 낮게 나타났다. 특히 왕벚나무는 q10 관측비율이 0.132로 목표 분위보다 높게 나타나 다른 수종보다 하위 분위의 재현 정도가 상대적으로 낮았다. 교차검증에서 산출된 q10과 q90의 합산 pinball loss는 수종별로 0.903–2.839의 범위로 나타났다.

표 7. 수종별 최종모형의 교차검증 성능과 분위회귀선 교차 여부

| 수종 | 모형 | q10 핀볼 로스 | q90 핀볼 로스 | 합산 핀볼 로스 | q10 관측비율 | q90 관측비율 | 중앙 80% 포괄률 | 분위선 교차 |
|---|---|---|---|---|---|---|---|---|
| 가래나무 | 선형 | 1.193 | 1.646 | 2.839 | 0.106 | 0.897 | 0.79 | 없음 |
| 가죽나무 | 이차 | 0.498 | 0.847 | 1.345 | 0.112 | 0.906 | 0.795 | 없음 |
| 갈참나무 | 이차 | 0.693 | 1.04 | 1.733 | 0.101 | 0.9 | 0.799 | 없음 |
| 감나무 | 선형 | 0.553 | 0.966 | 1.518 | 0.114 | 0.886 | 0.772 | 없음 |
| 고로쇠나무 | 선형 | 0.843 | 1.491 | 2.334 | 0.1 | 0.899 | 0.799 | 없음 |
| 느티나무 | 선형 | 0.818 | 1.405 | 2.223 | 0.102 | 0.901 | 0.799 | 없음 |
| 단풍나무 | 선형 | 0.529 | 0.936 | 1.465 | 0.102 | 0.894 | 0.791 | 없음 |
| 당단풍나무 | 선형 | 0.448 | 0.73 | 1.177 | 0.1 | 0.898 | 0.798 | 없음 |
| 마가목 | 선형 | 0.684 | 1.34 | 2.024 | 0.104 | 0.896 | 0.793 | 없음 |
| 밤나무 | 선형 | 0.868 | 1.467 | 2.335 | 0.101 | 0.899 | 0.798 | 없음 |
| 백합나무 | 선형 | 0.444 | 0.497 | 0.941 | 0.117 | 0.891 | 0.774 | 없음 |
| 버드나무 | 선형 | 0.886 | 1.582 | 2.468 | 0.1 | 0.895 | 0.795 | 없음 |
| 뽕나무 | 선형 | 0.563 | 0.991 | 1.554 | 0.111 | 0.885 | 0.774 | 없음 |
| 산딸나무 | 선형 | 0.57 | 0.915 | 1.485 | 0.101 | 0.892 | 0.791 | 없음 |
| 상수리나무 | 선형 | 0.834 | 1.243 | 2.077 | 0.101 | 0.899 | 0.798 | 없음 |
| 왕벚나무 | 선형 | 1.301 | 1.474 | 2.775 | 0.132 | 0.891 | 0.759 | 없음 |
| 은사시나무 | 선형 | 0.812 | 1.367 | 2.179 | 0.108 | 0.899 | 0.791 | 없음 |
| 자귀나무 | 선형 | 0.585 | 0.934 | 1.52 | 0.104 | 0.899 | 0.796 | 없음 |
| 자작나무 | 선형 | 0.424 | 0.622 | 1.046 | 0.095 | 0.902 | 0.806 | 없음 |
| 쪽동백나무 | 선형 | 0.424 | 0.637 | 1.061 | 0.105 | 0.894 | 0.789 | 없음 |
| 층층나무 | 선형 | 0.914 | 1.329 | 2.243 | 0.101 | 0.9 | 0.799 | 없음 |
| 팽나무 | 선형 | 0.859 | 1.326 | 2.185 | 0.103 | 0.897 | 0.794 | 없음 |
| 피나무 | 선형 | 0.892 | 1.732 | 2.624 | 0.103 | 0.9 | 0.796 | 없음 |
| 함박꽃나무 | 선형 | 0.333 | 0.57 | 0.903 | 0.108 | 0.888 | 0.78 | 없음 |
| 황벽나무 | 선형 | 0.779 | 1.141 | 1.919 | 0.098 | 0.902 | 0.804 | 없음 |

다만 pinball loss의 크기는 수종별 흉고직경의 분포 범위와 변동성에 영향을 받으므로, 서로 다른 수종 사이의 절대적인 모형 우수성을 판정하는 기준으로 사용하지 않았다. 해당 값은 동일 수종 내에서 후보모형을 비교하고 최종모형을 선정하기 위한 상대적 예측오차로 해석하였다.

마지막으로 수종별로 관측된 실측수고의 최솟값과 최댓값 사이에 500개의 평가 지점을 생성하여 q10과 q90 분위회귀선의 교차 여부를 확인하였다. 분석대상 25개 수종 모두에서 q10 예측값이 q90 예측값보다 높아지는 지점은 확인되지 않았다. 따라서 최종모형은 각 수종에서 관측된 실측수고 범위 내에서 조건부 하위 분위와 상위 분위의 순서를 일관되게 유지하였다.

### 3.6 민감도 분석 결과

분석자료 선정 기준이 최종 분석대상에 미치는 영향을 확인하기 위해 고유 표본점 30개 이상 기준을 유지하면서 최소 관측값 기준을 50개, 75개, 100개 및 150개로 변경하였다. 모든 기준에서 분석대상은 동일한 25개 수종으로 나타났다. 이는 고유 표본점 기준을 충족한 수종 가운데 관측값이 가장 적은 왕벚나무도 174개의 유효 관측값을 보유하고 있었기 때문이다. 따라서 최소 관측값 기준을 50–150개 범위에서 변경하더라도 분석대상 수종의 구성은 달라지지 않았다. 해당 결과는 설정한 최소 관측값 기준의 변화에 대해 분석대상 수종 구성이 동일하게 유지되었음을 의미한다. 다만 개별 수종의 표본 수를 단계적으로 변화시키면서 q10 및 q90 회귀계수나 예측선의 안정성을 평가한 결과는 아니다.

반복 관측 가능성의 영향을 확인하기 위해 표본점번호와 개체목 번호가 동일한 반복 개체 후보에서 최초 관측값 한 건만 남기고 최종모형을 다시 추정하였다. 이에 따라 분석 관측값은 전체 통합자료의 64,072개에서 56,651개로 감소하였다. 그러나 25개 수종에서 산출된 q10과 q90의 1차항 추정치 50개는 모두 전체 통합자료와 동일한 양의 방향을 유지하였다. 전체 통합자료 대비 1차항 상대차이 절댓값의 중앙값은 1.61%로 나타났다. 다만 함박꽃나무 q90의 1차항은 전체 통합자료의 0.571에서 반복 개체 후보의 최초 관측값만 사용한 경우 0.784로 증가하여 가장 큰 상대차이인 37.25%를 보였다(표 8 참조).

조사차기에 따른 결과의 변화를 확인하기 위해 제5차, 제6차 및 제7차 자료를 각각 분리하여 전체 통합자료에서 선정된 것과 동일한 형태의 모형을 다시 추정하였다. 조사차기별 관측값이 50개 이상인 수종을 대상으로 분석한 결과, 각 조사차기에서 23개 수종의 q10 및 q90 계수 46개가 산출되었다. 모든 조사차기에서 1차항은 전체 통합자료와 동일한 양의 방향으로 나타났다(표 8 참조).

전체 통합자료 대비 1차항 상대차이 절댓값의 중앙값은 제5차에서 13.78%, 제6차에서 5.16%, 제7차에서 4.52%로 나타났다. 조사차기를 분리한 경우에도 실측수고가 큰 개체일수록 흉고직경의 조건부 분위값이 크게 나타나는 관계의 방향은 유지되었다. 다만 표본 수가 상대적으로 적은 일부 수종에서는 계수 크기의 차이가 컸다. 특히 함박꽃나무 q90의 1차항은 제6차에서 전체 통합자료보다 84.21% 크게 나타났으며, 제7차에서는 0에 가까운 값으로 추정되었다(표 8 참조).

종합하면, 최소 관측값 기준의 변화에 대해서는 분석대상 수종의 구성이 동일하게 유지되었다. 이 결과는 수종 선정기준의 민감도에 관한 것이며, 개별 수종에서 표본 수 변화에 따른 회귀계수나 예측선의 안정성을 의미하지는 않는다. 이와 별도로 반복 개체 후보의 최초 관측값만 사용한 분석과 조사차기별 분석에서는 추정 가능한 1차항의 방향이 전체 통합자료와 동일하게 나타났다.

## 4. 고찰

본 연구의 q10과 q90 분위회귀선은 같은 실측수고를 가진 수목이라도 흉고직경이 하나의 값으로 정해지지 않고

표 8. 반복 관측 처리 및 조사차기별 민감도 분석 결과

| 자료 구분 | 분석 수종 수 | 비교 계수 수 | 1차항 방향 일치율(%) | 상대차이 절댓값 중앙값(%) | 상대차이 절댓값 최댓값(%) |
|---|---|---|---|---|---|
| 반복 개체 후보별 최초 관측 | 25 | 50 | 100 | 1.61 | 37.25 |
| 제5차 | 23 | 46 | 100 | 13.78 | 53.33 |
| 제6차 | 23 | 46 | 100 | 5.16 | 84.21 |
| 제7차 | 23 | 46 | 100 | 4.52 | 100.00 |

일정한 범위로 나타난다는 점을 보여준다. 평균회귀선은 대표적인 흉고직경 값을 제시할 수 있지만, 상대적으로 가는 개체와 굵은 개체가 어떻게 분포하는지는 충분히 보여주기 어렵다. 반면 q10과 q90은 같은 수고 조건에서 흉고직경의 하위와 상위 위치를 구분하여 제시할 수 있으므로, 수종별 구조적 차이를 더 구체적으로 설명할 수 있다.

기존 연구는 주로 흉고직경과 임분 및 입지변수를 이용하여 평균 수고 또는 수고-흉고직경을 예측하였다. 경쟁, 입지 및 임분발달 변수를 포함한 비선형 혼합효과모형은 수고-흉고직경의 평균적 변화를 설명하는 데 활용되어 왔다(Sharma et al., 2019). 한편 분위회귀는 평균회귀만으로 충분히 표현하기 어려운 경쟁상태별 조건부 분포의 차이를 파악하는 데 활용될 수 있다(Zhang et al., 2020). 따라서 본 연구는 이러한 분위회귀의 장점을 국내 여러 낙엽활엽수종에 적용하여, 평균적인 관계가 아니라 실측수고를 조건으로 한 흉고직경의 하위 및 상위 분위범위를 제시했다는 점에서 기존 연구와 차별된다.

분석대상 수종에서 q90의 1차항이 q10보다 크게 나타난 결과는 개체 간 실측수고 차이에 대응하는 흉고직경 분위값의 차이가 하위 분위보다 상위 분위에서 더 크게 나타났음을 의미한다. 이는 본 연구의 관측자료에서 확인된 조건부 분포의 통계적 특징이며, 특정한 생태적 요인에 의해 발생한 결과를 의미하지는 않는다. 동일한 수고를 가진 수목 사이의 흉고직경 차이는 수령, 경쟁상태, 임분밀도 및 입지조건 등과 관련되었을 가능성이 있으나, 본 연구에서는 이러한 요인들을 설명변수로 포함하여 각각의 영향을 직접 검증하지 않았으므로 q10과 q90의 차이를 특정 요인의 영향으로 단정할 수 없다. 따라서 본 연구에서 확인된 분위별 차이는 실측수고를 조건으로 한 흉고직경 분포의 구조적 차이로 한정하여 해석하였으며, 관련 생태적 요인의 영향은 해당 변수들을 포함한 후속 연구에서 검증할 필요가 있다.

대부분의 수종에서 선형모형이 선정되고 일부 수종에서만 이차모형이 선정된 결과는 복잡한 모형이 항상 우수한 예측성능을 보장하지 않는다는 점을 보여준다. Cubic B-spline은 자료의 세부적인 굴곡을 유연하게 표현할 수 있지만 특정 자료에 지나치게 적합될 가능성이 있다. 또한 본 연구에서는 표본점 단위 교차검증과 1 표준오차 규칙을 적용하여, 복잡한 모형의 성능개선이 충분히 확인될 때만 비선형모형을 선택하였다. 이는 모형의 복잡성보다 새로운 표본점에서 조건부 분위값을 안정적으로 재현하고 결과를 쉽게 해석할 수 있는지를 고려한 것이다.

표본점 단위 교차검증에서 관측비율과 중앙 80% 포괄률이 목표값에 근접했으며, 분위회귀선의 교차도 나타나지 않았다. 이는 각 교차검증에서 모형 추정에 포함되지 않은 NFI 표본점의 관측값에 대해 목표 분위와 중앙 80% 범위를 대체로 재현하였음을 보여준다. 다만 이는 NFI 자료 내부의 표본점 단위 교차검증의 결과이므로, 조사대상과 생육환경이 다른 도시숲이나 가로수에 대한 외부 타당성을 의미하지 않는다. 또한 표본점 단위 군집 부트스트랩을 적용하여, 같은 표본점에 속한 수목들이 공유하는 환경과 관측값의 유사성을 고려한 상태에서 회귀계수의 불확실성을 평가하였다.

본 연구는 국내 대규모 NFI 자료에 수종별 분위회귀 기반 조건부 분포 분석체계를 적용하고, 내부 검증성능과 회귀계수의 불확실성을 제시한 방법론적·실증적 연구이다. 구체적으로 25개 수종의 q10과 q90 회귀식 및 적용 가능한 실측수고 범위를 제시하여, 평균적인 관계만으로는 확인하기 어려운 흉고직경 조건부 분포의 하위와 상위 위치를 정량화하였다. 또한 표본점 군집구조를 반영한 교차검증과 군집 부트스트랩을 통해 예측성능과 회귀계수의 불확실성을 함께 평가하였다는 데 학술적 의의가 있다.

다만 실측수고가 기록된 개체목만 분석에 포함하였으므로 결과가 NFI 전체 조사목을 완전히 대표한다고 보기 어렵다. 설명변수도 실측수고 하나로 한정하여 수령, 경쟁상태, 임분밀도, 지역 및 입지조건의 영향을 분리하여 검증하지 못하였다. 최소 관측값 기준을 50-150개로 변경한 민감도 분석 역시 분석대상 수종 구성의 변화를 평가한 것이다. 모든 기준에서 동일한 25개 수종이 선정되었으나, 이는 개별 수종의 표본 수 변화에 따른 q10 및 q90 회귀계수나 예측선의 안정성을 직접 검증한 결과로 해석할 수 없다.

조경적 측면에서 본 연구의 결과는 향후 도시수목의 독립적인 실측자료를 이용한 외부검증을 설계하고, 도시수목 고유의 조건부 분위범위를 구축하기 위한 산림수목 기반 참조정보로 의미가 있다. 현 단계에서 제시한 회귀식을 도시수목의 생육상태 판정이나 정밀조사 대상의 우선순위 결정에 직접 적용하는 것은 적절하지 않다.

향후 연구에서는 도시숲과 가로수의 독립자료를 이용하여 q10과 q90의 관측비율 및 중앙 80% 포괄률을 평가하고, 수령, 경쟁상태, 토양용적, 불투수면, 관수 및 전정 이력에 따른 차이를 검토할 필요가 있다. 또한 동일 개체목의 반복측정자료를 활용하여, 본 연구의 여러 조사차수 통합자료에서 확인된 개체 간 수고-흉고직경의 구조적 관계와 동일 개체목의 시간에 따른 수고 및 흉고직경 변화를 구분하여 검토할 필요가 있다. 도시수목 독립자료를 이용한 외부검증을 통해 적용 가능성이 확인된다면, 본 연구의 분석체계는 수종과 수고 조건을 고려하여 구조적으로 특이한 개체를 탐색하고 추가 현장조사의 필요성을 검토하는 보조자료로 확장될 수 있을 것이다.

## 5. 결론

본 연구는 국가산림자원조사 자료를 활용하여 낙엽활엽수종별 실측수고에 따른 흉고직경의 조건부 하위 및 상위 분위범위를 제시하고자 수행하였다. 흉고직경을 포함한 수고곡선식으로 산정되는 추정수고를 다시 흉고직경의 설명변수로 사용할 경우 발생할 수 있는 순환논리를 해소하기 위해 추정수고를 제외하고, 현장에서 측정된 실측수고와 흉고직경이 모두 유효한 자료만을 사용하였다. 최종적으로 25개 수종, 64,072개 관측값을 대상으로 q10 및 q90 분위회귀모형을 구축하였다.

후보모형의 표본점 단위 교차검증 결과, 가죽나무와 갈참나무에서는 이차모형이, 나머지 23개 수종에서는 선형모형이 선정되었다. 대부분의 수종에서 단순한 선형모형이 선택된 것은 복잡한 비선형모형이 항상 우수한 검증성능을 보장하지 않으며, 수종별 자료에서 확인되는 예측성능과 해석 가능성을 함께 고려해 함수형태를 결정할 필요가 있음을 보여준다. 또한 모든 수종에서 q90의 1차항이 q10보다 크게 나타나, 개체 간 실측수고 차이에 대응하는 흉고직경 분위값의 차이가 조건부 분포의 하위보다 상위에서 상대적으로 크게 나타났다.

최종모형은 표본점 단위 교차검증에서 목표 분위와 중앙 80% 범위를 전반적으로 재현하였으며, 관측된 실측수고 범위에서는 q10과 q90 분위회귀선의 교차가 나타나지 않았다. 표본점 단위 군집 부트스트랩을 통해 동일 표본점의 관측값이 공유하는 자료구조를 유지하면서 회귀계수의 불확실성도 제시하였다. 이에 따라 본 연구는 Pseudo $R^2$와 같은 하나의 표본 내 적합도 지표에 의존하지 않고, 검증자료의 예측오차, 관측비율, 포괄률, 계수의 불확실성과 분위선의 논리적 일관성을 함께 평가하였다.

본 연구의 학술적 의의는 독립적으로 측정된 실측수고를 이용하여 순환논리를 해소하고, 평균적인 관계만으로는 확인하기 어려운 흉고직경 조건부 분포의 하위와 상위 위치를 25개 수종별로 제시한 데 있다. 조경 분야에서 본 연구의 결과는 향후 도시숲과 가로수의 독립적인 실측자료를 이용한 외부검증을 설계하고, 도시수목 고유의 조건부 분위범위를 구축하기 위한 산림수목 기반 참조정보로 활용될 수 있을 것이며, 향후 외부검증을 통해 도시수목에 대한 적용 가능성이 확인된다면, 조건부 분포에서 상대적으로 특이한 개체를 파악하고 수관상태, 병해충, 토양환경 및 관리이력에 대한 추가 조사의 필요성을 검토하는 기초자료로 활용될 수 있을 것이다.

다만 본 연구는 산림수목을 조사한 NFI 자료를 이용하였으므로, 그 결과를 도시수목의 생육상태 판단에 활용할 수 있는지는 별도의 검증이 필요하다. 또한 본 연구의 q10과 q90은 같은 수종에서 동일한 실측수고를 가진 수목들의 흉고직경 분포에서 하위와 상위 위치를 나타내며, 개별 수목의 미래 생장량이나 생물학적으로 가능한 최대 흉고직경을 의미하지 않는다. 이러한 해석 범위와 함께 자료의 수고 범위도 고려하여, 수종별 회귀식은 분석자료에서 확인된 실측수고 범위 내에서 해석해야 한다. 향후에는 도시숲과 가로수의 수고, 흉고직경, 수령, 입지 및 관리조건을 포함한 독립자료를 구축하여 도시수목에 대한 외부 타당성을 검증할 필요가 있다. 이와 함께 동일 개체목의 반복측정자료를 활용하여, 본 연구에서 확인한 개체 간 구조적 관계와 동일 개체목의 시간에 따른 생장 변화를 구분하여 분석할 필요가 있다.

### References


1. 강진택(2019) 국가산림자원조사 자료를 이용한 최저-최고 흉고직경 추정식 개발. 농업생명과학연구 53(6): 23-33.
2. 산림청(2022a) 제5차 국가산림자원조사 마이크로데이터(2006-2010) [데이터세트]. 산림임업통계플랫폼.
3. 산림청(2022b) 제6차 국가산림자원조사 마이크로데이터(2011-2015) [데이터세트]. 산림임업통계플랫폼.
4. 산림청(2022c) 제7차 국가산림자원조사 마이크로데이터(2016-2020) [데이터세트]. 산림임업통계플랫폼.
5. Almas, A. D. and T. M. Conway(2016) The role of native species in urban forest planning and practice: A case study of Carolinian Canada. Urban Forestry & Urban Greening 17: 54-62.
6. Allen, C. D., D. D. Breshears and N. G. McDowell(2015) On underestimation of global vulnerability to tree mortality and forest die-off from hotter drought in the Anthropocene. Ecosphere 6(8): 1-55.
7. Amini Parsa, V., E. Salehi, A. R. Yavari and P. M. van Bodegom(2019) Evaluating the potential contribution of urban ecosystem service to climate change mitigation. Urban Ecosystems 22(5): 989-1006.
8. Bohora, S. B. and Q. V. Cao(2014) Prediction of tree diameter growth using quantile regression and mixed-effects models. Forest Ecology and Management 319: 62-66.
9. Condés, S. and H. Sterba(2005) Derivation of compatible crown width equations for some important tree

species of Spain. Forest Ecology and Management 217(2-3): 203-218.
10. Esperon-Rodriguez, M., M. G. Tjoelker, J. Lenoir, J. B. Baumgartner, L. J. Beaumont, D. A. Nipperess, S. A. Power, B. Richard, P. D. Rymer and R. V. Gallagher(2022) Climate change increases global risk to urban forests. Nature Climate Change 12(10): 950-955.
11. Grima, N., W. Corcoran, C. Hill-James, B. Langton, H. Sommer and B. Fisher(2020) The importance of urban natural areas and urban ecosystem services during the COVID-19 pandemic. PLOS ONE 15(12): e0243344.
12. Huang, K.(2022) Urban forests facing climate risks. Nature Climate Change 12(10): 893-894.
13. Ige, P. and O. Komolafe(2022) Tree slenderness coefficient models for biodiversity conservation in International Institute of Tropical Agriculture Forest Ibadan, Nigeria. Tanzania Journal of Forestry and Nature Conservation 91(1): 20-31.
14. Kang, J., C. Ko, S. J. Lee, J. S. Yim, G. H. Moon and S. H. Lee(2021) Relationship of H/D and crown ratio and tree growth for Chamaecyparis obtusa and Cryptomeria japonica in Korea. Forest Science and Technology 17(3): 101-109.
15. Koenker, R.(2017) Quantile regression: 40 Years On. Annual Review of Economics 9: 155-176.
16. Marchin, R. M., M. Esperon-Rodriguez, M. G. Tjoelker and D. S. Ellsworth(2025) Understanding urban tree heat and drought stress by tracking growth and recovery following an extreme year. Landscape and Urban Planning 261: 105394.
17. Muscas, D., M. Fornaciari, C. Proietti, L. Ruga and F. Orlandi(2023) Tree growth rate under urban limiting conditions. European Journal of Forest Research 142(6): 1423-1437.
18. Pirselimoğlu Batman, Z. and E. Ender Altay(2025) Recreational performance evaluation of urban forests: Spatial, socio-cultural, and public health-related perspectives. International Journal of Environmental Research and Public Health 22(9): 1401.
19. Pretzsch, H., P. Biber, E. Uhl, J. Dahlhausen, G. Schütze, D. Perkins, T. Rötzer, J. Caldentey, T. Koike, T. van Con, A. Chavanne, B. du Toit, K. Foster and B. Lefer(2017) Climate change accelerates growth of urban trees in metropolises worldwide. Scientific Reports 7(1): 15403.
20. Rahim, N. H., S. A. Razak, X. Chang, F. C. Saun, M. N. Khan, S. N. Hamzah, F. Rohman, B. Ali, A. Kaplan, M. Iqbal, M. R. Bozhuyuk, and S. Ercisli(2024) Ecosystem services by urban forest (UF) towards climate change adaptation: A review. Polish Journal of Environmental Studies 33(4): 3503-3513.
21. Raptis, D. I., V. Kazana, S. Kechagioglou, A. Kazaklis, C. Stamatiou, D. Papadopoulou and T. Tsitsoni(2022) Nonlinear quantile mixed-effects models for prediction of the maximum crown width of *Fagus sylvatica* L., *Pinus nigra* Arn. and *Pinus brutia* Ten. Forests 13(4).
22. Scarmana, G. and K. McDougall(2023) Assessing the relationship between tree dimensions in an urban environment. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 48: 437-442.
23. Sharma, R. P., Z. Vacek, S. Vacek and M. Kučera(2019) A nonlinear mixed-effects height-to-diameter ratio model for several tree species based on Czech National Forest Inventory Data. Forests 10(1): 70.
24. Sumida, A., T. Miyaura and H. Torii(2013) Relationships of tree height and diameter at breast height revisited: Analyses of stem growth using 20-year data of an even-aged Chamaecyparis obtusa stand. Tree Physiology 33(1): 106-118.
25. Zhang, B., S. Sajjad, K. Chen, L. Zhou, Y. Zhang, K. K. Yong and Y. Sun(2020) Predicting tree height-diameter relationship from relative competition levels using quantile regression models for Chinese Fir (*Cunninghamia lanceolata*) in Fujian Province, China. Forests 11(2).