

多地点・多時期データによる樹高曲線のモデル選択

Height-diameter model selection using multi-site and multi-temporal data

細田和男^{*1}・西園朋広^{*1}・山田祐亮^{*1}・高橋正義^{*1}・齋藤英樹^{*1}・佐野真琴^{*1}Kazuo HOSODA^{*1}, Tomohiro NISHIZONO^{*1}, Yusuke YAMADA^{*1}, Masayoshi TAKAHASHI^{*1}, Hideki SAITO^{*1} and Makoto SANO^{*1}

* 1 森林総合研究所 森林管理研究領域

Department of Forest Management, Forestry and Forest Products Research Institute

要旨: 関東森林管理局および中部森林管理局の国有林に所在するスギ、ヒノキまたはカラマツの人工林収穫試験地毎木調査データ 87 試験区 482 回分を用い、非線形最小二乗法あるいは通常の線形回帰によって、Henricksen 式、Stoffels 式、Näslund 式、拡張相対成長式および修正ワイブル式の 5 種類の樹高曲線モデルへのあてはめを行った。バイアス、平均二乗誤差平方根、赤池の情報量基準に加え、回帰係数の妥当性を考慮して評価した結果、全体としては Näslund 式のあてはまりが良好であるという通説の妥当性が確認された。また同時に、樹高曲線モデルによる精度の差はわずかであることも再認識された。非線形最小二乗法ではなく、直線変換後の線形回帰で代用した場合の樹高の推定バイアスは、実用上それほど大きいものではないことが明らかになった。もともと通常の線形回帰であてはめることのできる Henricksen 式は、直線変換した Näslund 式に匹敵する精度を示したことから、非線形最小二乗法が実行できない場合において、第一の選択肢になるものと考えられた。

キーワード: ヘンリクセン式・ネズルンド式・相対成長式・修正ワイブル式

Abstract: As height-diameter model, Henricksen equation, Stoffels equation, Näslund equation, extended allometric equation and modified Weibull equation were tested using multi-site and multi-temporal data collected from 87 permanent experimental plots in national forests, Kanto and Chubu Regional Forest Offices. The fitness of five models was evaluated by mean error, RMSE and Akaike's information criterion. The advantage of Näslund equation was acknowledged similar to previous study. And there was a slight difference of accuracy between five height-diameter models. Henricksen equation is a second accurate model after Näslund equation and is calculable with ordinary least-square method. Thus Henricksen equation is alternative option if nonlinear least-squares method is not available.

Key-word: Henricksen equation, Näslund equation, allometric equation, modified Weibull equation

はじめに

能率がよく高精度な測高器が比較的安価に入手できるようになって久しい。しかし実務上の毎木調査においては全ての立木の樹高を実測することは稀であり、樹高を実測しなかった立木の樹高は、樹高曲線を介して胸高直径から推定することが多い。測樹学分野では、過去さまざまな樹高曲線のモデルが用いられ、例えば西沢(5)には 19 種類、説明変数として林齢を追加したモデルを含めると実に 25 種類もの樹高曲線式が紹介されている。これらさまざまな樹高曲線モデルの中では、Näslund 式のあてはまりが相対的にやや良好であることが知られている(3, 6, 4)。しかしながら、前二者の先行研究は、コンピュータが普及する以前の時代のもので、本来は曲

線であるモデルを直線変換する方法であてはめが行われた可能性が高い。また三者とも特定の地域の、比較的少数の事例データから得られた結果である。さらに、比較対象になった樹高曲線モデルには、森林生態学分野で多用される拡張相対成長式が含まれていないなど、再検討の余地が残されているといえる。

本研究では、多地点で多時期に収集された同齢単純林の毎木調査データを用い、非線形最小二乗法あるいは直線変換によって、いくつかの代表的な樹高曲線モデルへのあてはめを行う。この結果から、RMSE や AIC などモデル評価基準として、各樹高曲線モデルやあてはめ方法の妥当性を比較検討することを目的とする。

資料と方法

1. 毎木調査データ 樹高曲線のあてはめには、関東森林管理局および中部森林管理局の国有林に所在するスギ、ヒノキまたはカラマツの人工林収穫試験地(1)の毎木調査データを用いた。収穫試験地では原則として毎回全ての立木の樹高が実測されているが、樹高が抽出測定にとどまった調査回もあり、樹高の実測本数率が95%以上の調査回を解析対象とした。目的樹種以外の侵入樹種や枯死木は除外した。一方、試験地内に散見される梢端折れ、幹折れ被害木や被圧木は、そのような被害木や劣勢木が多少とも混在するのが現実の人工林であることから、生立木であれば除外せずに解析に供した。結果として採用したデータセットは87試験区の482回分で、調査本数の合計は130,171本である。採用した試験区の面積は0.075~0.4ha、平均0.18ha、調査時の林齢は5~102年、平均43年、各調査回の調査本数は61~939本、平均270本であった。

2. 樹高曲線モデル 前述の482回分のデータセットに対し、以下の5種類の代表的な樹高曲線モデルをあてはめた。

$$H = a + b \log D \quad (1)$$

$$H = a D^b \quad (2)$$

$$H = (D / (a + b D))^2 \quad (3)$$

$$H = 1 / (1 / (a D^b) + 1 / c) \quad (4)$$

$$H = a (1 - \exp(-b D^c)) \quad (5)$$

ここで H は樹高(m)、 D は胸高直径(cm)、 $a \sim c$ は回帰係数である。

モデル(1)~(3)はそれぞれ Henricksen 式、Stoffels 式、Näslund 式と呼ばれ、(3)、(4)、(6)がいずれも比較検討の対象とした、測樹学分野での代表的な樹高曲線式である。モデル(2)は単純相対成長式でもある。モデル(4)は拡張相対成長式と呼ばれ、森林生態学分野で多用されている。モデル(5)は修正ワイブル式(8)と呼ばれる成長曲線の一種であるが、樹高曲線としても用いられる(2)。このモデル(5)について Thomas(7)は Von Bertalanffy の式と密接な関係があるモデルとし、モデル(4)との比較を行っている。モデル(4)の係数 c およびモデル(5)の係数 a は樹高の上限を表す係数である。

3. 樹高曲線のあてはめ方法 比較する樹高曲線はいずれも曲線であり、モデル(1)以外は通常の線形回帰ではあてはめることができない。モデル(2)~(5)のあてはめは、統計解析ソフトウェア R の minpack.lm パッケージに含まれる nls.lm 関数を用い、Levenberg-Marquard 法による非線形最小二乗法で実行した。また比較のため、モデル

(2)と(3)は以下に示す一次式それぞれ(2)'と(3)'に変換し、通常の線形回帰を行う旧来の方法も同時に行った

$$\log H = \log a + b \log D \quad (2)'$$

$$D / \sqrt{H} = a + b D \quad (3)'$$

ここで各記号の意味は、式(1)~(5)と同じである。

3. モデル評価基準 得られた樹高曲線の評価基準として、以下の(6)~(8)式でそれぞれ定義される平均誤差 Bias、平均二乗誤差平方根 RMSE、赤池の情報量基準 AIC を用いた。

$$\text{Bias} = \sum(\hat{h} - h) / n \quad (6)$$

$$\text{RMSE} = \sqrt{\sum(h - \hat{h})^2 / n} \quad (7)$$

$$\text{AIC} = n \log(\sum(h - \hat{h})^2 / n) + 2p \quad (8)$$

ここで、 h は実測樹高、 \hat{h} は推定樹高、 n は本数、 p はモデルのパラメーター数である。

Bias が正の場合は推定樹高が過大推定、逆に負の場合は過小推定であることを表す。RMSE は小さいほど推定精度が高いことを示す。AIC はパラメーター数を加味したうえで、相対的に小さいほど良好なモデルであることを意味する。なお、モデル(2)'と(3)'については、それぞれ得られた回帰係数を直線変換前のモデル(2)および(3)に代入し、元の単位での樹高の推定誤差を算出した。

結果と考察

1. あてはめ方法による違い 482回分のデータセットに対し、5種類の代表的な樹高曲線をあてはめたときの Bias および RMSE を表-1に示す。

まず、非線形最小二乗法と直線変換による線形回帰による違い、すなわちモデル(2)と(2)'、モデル(3)と(3)'の相違に着目する。RMSE の平均は、モデル(2)と(3)ともに約1.1mであり、それぞれ対応する直線変換モデル(2)'や(3)'との間にも大きな違いはなかった。

一方 Bias は、モデル(2)と(3)がそれぞれ0.002m、-0.001mと極めてゼロに近いのに対し、モデル(2)'と(3)'はそれぞれ-0.041m、-0.062mと、やや過小推定の傾向を示していた。これは直線変換に伴う偏りであり、当然ながら非線形最小二乗法によるほうが、データに対して、よりあてはまりのよい樹高曲線が得られるといえる。もっとも、この偏りは今回検討したデータの範囲では、平均で-0.1m以内、最悪のケースでもモデル(2)'で-0.232m、モデル(3)'で-0.376mであった。したがって、樹高の測定値自体の不確実性を考慮すれば、とりわけ実務上の森林調査においては大きな問題ではない、という解釈もありうる。

2. 樹高直線モデルによる違い 次に通常の線形回帰であてはめ可能なモデル(1), および非線形最小二乗法であてはめたモデル(2)~(5)の差異に着目する(表-1)。モデル(1)は通常の線形回帰で求めており, Bias はゼロである。またモデル(2)~(5)の Bias も極めてゼロに近かった。

RMSE の比較では, モデル(4)が 1.073m で最も小さく, 次いで(5), (3), (1), 最も RMSE が大きいのはモデル(2)で 1.100m であった。モデル(4)と(2)の差は 0.027m に過ぎなかった。高田(6)は樹高曲線モデルによる精度の差は微小であること, また梶原(3)は Henricksen 式と Näslund 式の精度の差は極めて小さいことを指摘している。本研究の結果は, 多地点・多時期の豊富なデータセットから, これらの報告を追認したものといえる。

もう一つの評価基準である AIC については, その平均値を比較することには意味がないので, 482 の各データセットにおいて, AIC が小さいほうから各モデルに 1~5 位の順位付けを行い, その 482 回分の平均順位を表-2 に示した。全体としては, 平均順位が高いのはモデル(3)~(5), 次いでモデル(1), 平均順位が最も低いのはモデル(2)であった。前述のとおり, RMSE の比較ではモデル(4)や(5)がやや優位であったが, パラメーター数を加味した AIC ではモデル(3)~(5)がほぼ同等であると評価された。

さらに, 樹種や平均樹高クラスによって区分して比較すると(表-2), 樹種別では全体と同じ傾向であった。平均樹高クラス別では, 10m 未満のクラスではモデル(3)よりもモデル(4)や(5)が優位, スギの 30m 以上やヒノキ・カラマツの 20m 以上では, モデル(4)や(5)よりもモデル(1)や(3)のほうが優位といった特徴がみられた。

3. 回帰係数の妥当性 データセット 482 回分のうち 12 例, 本数では 26 本に過ぎないが, モデル(3)は小径木に負の樹高推定値を与えるケースがあった。また, 前述の通り, モデル(4)と(5)には最大樹高を示すパラメーターがあるが, 非線形最小二乗法によるあてはめの結果, そのパラメーターが非現実的な値に収束するケースが散見された。すなわち, モデル(4)の回帰係数 c が 100 以上であったのが 59 例, モデル(5)の回帰係数 a が 100 以上であったのが 34 例あった。特にモデル(4)は全体の 10% 以上において回帰係数 c が 50 以上であり, 平均樹高が 10m 未満の範囲を中心に c が 10 万以上の値に収束したのが 29 例あった(図-1)。これは同齢単純林の場合, 若齢時は胸高直径と樹高との関係が, 上に凸の飽和曲線よりも直線関係に近いのではないかと推察される。

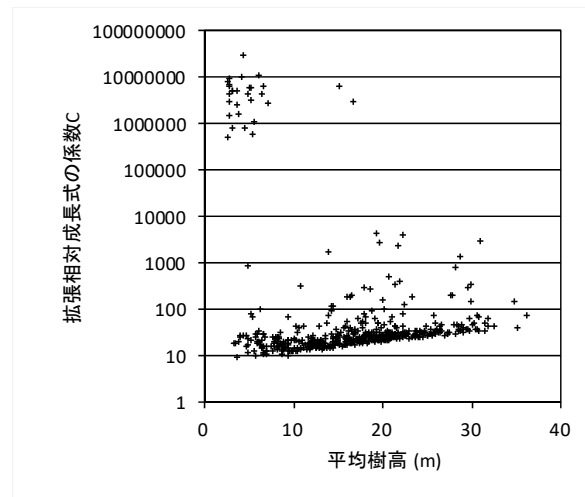


図-1. 平均樹高と拡張相対成長式の係数 c との関係
Fig.1 Relationship between mean height and parameter c of extended allometric equation

まとめ

前述のように, パラメーター数の違いを加味した場合, 5 種類の樹高曲線モデルの中では, 全体として Näslund 式, 拡張相対成長式および修正ワイブル式の 3 種類が高く評価された。平均樹高が 10m 以下の範囲では, AIC による Näslund 式の評価が低くなった一方, 拡張相対成長式および修正ワイブル式の最大樹高を表すパラメーターが非現実的な値に収束するケースがあった。また, 樹高の高い範囲では拡張相対成長式や修正ワイブル式よりも Näslund 式が優位であった。

以上の検討から, 多地点・多時期のデータセットと非線形最小二乗法の下でも, 全体的にみて Näslund 式のあてはまりが良好であるという通説の妥当性が確認された。また同時に, 高田(6)や梶原(3)が指摘した通り, 特に実用上, 樹高曲線モデルによる精度の差はわずかであることも再認識された。

非線形最小二乗法が実行できない環境において, 直線変換で代用した場合の推定バイアスは, 同じく実用上それほど大きいものではないことも明らかになった。表-1 のモデル(1)と(2)との違いに着目すると, RMSE に大差はなく, Bias は Henricksen 式のほうが小さい。もともと通常の線形回帰であてはめることのできる Henricksen 式は, 非線形最小二乗法が実行できない場合において, 第一の選択肢になるものと考えられた。

引用文献

- (1) 細田和男・家原敏郎・鷹尾元・西園朋広・高橋與明・石橋聡・高橋正義・古家直行・小谷英司・田中邦宏・

平田泰雅・光田靖・北原文章・近藤洋史・齋藤英樹・佐野真琴(2014)平成18~22年度に調査した収穫試験地等固定試験地の経年成長データ(収穫試験報告第25号). 森林総研報 13 : 225-254

(2) HUANG, S., TITUS, S. J. and WIENS, D. P. (1992) Comparison of nonlinear height-diameter functions for major Alberta tree species. Can. J. For. Res. 22 : 1297-1304

(3) 梶原幹弘(1966)樹高曲線の誤差に関する実験的考察. 京府大演報 10 : 1-14

(4) 國崎貴嗣・川村かの子・齋藤信人(1999)密度管理の違いがスギ人工同齢単純林の樹高曲線の形状に及ぼ

す影響. 岩大演報 30 : 97-108

(5) 西沢正久(1972)森林測定. 農林出版, 東京 : 348pp

(6) 高田和彦(1959)樹高曲線式と材積式との関係. 日林誌 41 : 275-276

(7) THOMAS, S. C. (1996) Asymptotic height as a predictor of growth and allometric characteristics in Malaysian rain forest trees, American J. Botany 83 : 556-566

(8) YANG, R. C., KOZAK, A. and SMITH, J. H. G. (1978) The potential of Weibull-type functions as a flexible growth curves. Can. J. For. Res. 8 : 424-431

表 - 1 . 各樹高曲線モデルにおける Bias と RMSE

Table 1 Biases and RMSEs in five height-diameter models

モデル	Bias(m)			RMSE(m)		
	平均	最大	最小	平均	最大	最小
(1)	0.000	0.000	0.000	1.090	3.252	0.041
(2)	0.002	0.032	-0.003	1.100	3.251	0.113
(3)	-0.001	0.016	-0.051	1.085	3.257	0.036
(4)	-0.001	0.006	-0.021	1.073	3.251	0.028
(5)	0.000	0.008	-0.008	1.074	3.251	0.029
(2)	-0.041	0.002	-0.232	1.107	3.254	0.119
(3)	-0.062	0.000	-0.376	1.088	3.264	0.036

(N=482)

表 - 2 . AIC のモデル別平均順位

Table 2 Average rank of Akaike's Information Criteria by five models

樹種	平均樹高 クラス	データ セット数	モデル				
			(1)	(2)	(3)	(4)	(5)
スギ	全体	173	3.2	3.7	2.7	2.5	2.8
	< 10m	45	4.4	3.0	3.6	2.0	2.0
	< 20m	58	2.7	4.0	2.6	2.6	3.1
	< 30m	57	3.0	4.2	2.2	2.6	2.9
	30m	13	1.9	2.9	2.5	3.7	3.9
ヒノキ	全体	181	3.1	3.8	2.7	2.7	2.8
	< 10m	42	3.5	3.1	4.0	2.2	2.3
	< 20m	100	3.1	4.3	2.2	2.6	2.7
	20m	39	2.7	3.3	2.4	3.2	3.3
カラマツ	全体	128	3.0	3.8	2.6	2.8	2.8
	< 10m	25	3.4	3.3	3.9	2.1	2.2
	< 20m	62	3.0	4.0	2.4	2.9	2.8
	< 30m	39	2.8	3.9	2.1	3.1	3.1
	30m	2	1.5	3.5	1.5	4.0	4.5
樹種計	クラス計	482	3.1	3.8	2.7	2.7	2.8