

## 混合モデルを用いた回帰式による成長経過の解析の一例

那須仁弥(森林総研林育セ)

要旨:成長経過は個体に対して時間経過とともに繰り返し測定されているので、得られたのは経時データとなる。これに回帰式を当てはめることによって、経時変化をモデル化することが出来る。本報告では北海道江別市にある森林総合研究所林木センター北海道育種場内のミズナラ産地別試験地の記録に対して回帰式を当てはめて解析した。実測値が推定された回帰式の95%信用区間内にあり、よく当てはまっていると判断できた。推定された回帰式の傾き、切片の比較により、成長経過に及ぼす複数の要因の比較や産地間、家系間などのグループ間の比較が出来た。これらのことから混合モデルを用いた回帰式は、成長経過の解析に適していると考えられた。

キーワード:成長経過, 混合モデル, 回帰式

## I はじめに

樹高や胸高直径などの成長経過は、個体に対して時間経過とともに繰り返し測定することが可能であり、得られたのは経時データとなる。調査期間を通しての時間経過に伴う変化をつかむためには、回帰式や曲線式を当てはめることによって、成長経過をモデル化することが出来る。しかし、産地-家系-集団などの階層をもつデータの場合、個体レベルでの回帰式の当てはめは個体変動には考慮できるが、集団全体の平均的な成長経過をとらえることが出来ない。一方、集団レベルで当てはめられた回帰式は個体変動をとらえることが出来ない。ここで個体ごとの回帰式が集団平均の周りで変動することを想定する。このとき、個体ごとの回帰式の回帰係数をランダム係数(変量効果)とする混合モデルでは集団の平均と各個体の成長曲線が同時に推定できる(2)。林木においても、このではランダム係数をつかった樹高成長の解析の報告がなされている(1, 3)。本報告ではミズナラの人工植栽試験地で植栽後26年までに得られた断片的な調査結果を用いて成長経過について、混合モデルを用いた回帰式の有効性について検討を行った。

## II 材料と方法

林木育種センター北海道育種場(北海道江別市)の場内に設定された第1ミズナラ産地別試験園に植栽されたミズナラ自然交配家系の樹高データを用いた。

第1試験園には1980から1981年にかけて道内12産地、64家系が植栽され、調査は7, 10, 20, 26年次に行われている。解析はデータを対数変換し、産地ごとまたは家系ごとの回帰式の回帰係数をランダム係数(変量効果)とする混合モデルを仮定すると、集団の平均と産地ごとまたは家系ごと

の回帰式が同時に推定できる(1)ので、以下の式を仮定した。

$$Y_{ijk} = \mu + S_j + \alpha_j T_{ijk} + F_k + \beta_k T_{ijk} + e_{ijk}$$

ここでは、 $Y_{ijk}$ は*i*年次の*j*番目の産地の*k*番目の産地内家系の*i*番目の個体、 $\mu$ は全体平均、 $T_i$ は年次、 $S_j$ は産地の切片、 $\alpha_j$ は年次への産地のランダム1次偏回帰係数、 $F_k$ は産地内家系のランダム切片、 $\beta_k$ は年次への産地内家系のランダム1次偏回帰係数、 $e_{ijk}$ は残差をしめす。

$$Y = X\beta + Z_1a + Z_2b + e$$

$Y$ :各調査年次での樹高の観測ベクトル、 $X$ :母数効果の計画行列、 $\beta$ :未知の母数効果(年次の効果)ベクトル、 $Z_1$ :変量効果の計画行列、 $a$ :未知の変量効果(産地の効果)のベクトルで平均0、 $b$ :未知の変量効果(産地内家系の効果)のベクトルで平均0、分散共分散行列  $G_2$ 、 $e$ :誤差のベクトルで平均0、分散共分散行列  $R$ とした。これらの解析にはR言語(3)で行った。

## III 結果と考察

解析では樹高を対数変換したので図には予測値を逆対数変換してplotした。

## 1. 推定された回帰式の当てはまり

実測値が推定された回帰式の95%信用区間内にあり、よく当てはまっていると判断できた(図-1)。

## 2. 回帰式の要因の分散の推定

年次と産地内家系共分散以外の要因は95%信用区間に0が含まれず、効果があると判断された。産地の切片、年次に対するランダム係数ともに産地内家系より大きく、効果の大きさは産地内家系より産地の方が大きいと考えられた(表-1)。

### 3.産地間の比較および産地別回帰式の推定

産地の年次へのランダム係数および切片を表-2に示す。図-2に足寄および陸別の産地別回帰式を示す。推定された産地別の回帰式にはランダム切片の大きさが足寄で-0.352、陸別で0.31とよく反映されていると思われた。これらのことから混合モデルを用いた回帰式の解析は成長経過に及ぼす複数の要因の比較、産地間、家系間などのグループ間の比較に有用であると考えられた。試験地の設定、管理に関わった歴代の北海道育種場の職員方々に深謝する。

### IV 引用文献

- (1) Chunkao Wang, Bengt Andersson, and Patrik Waldmann (2009) Genetic analysis of longitudinal height data using random regression. Can. J. For. Res. 39(10): 1939-1948

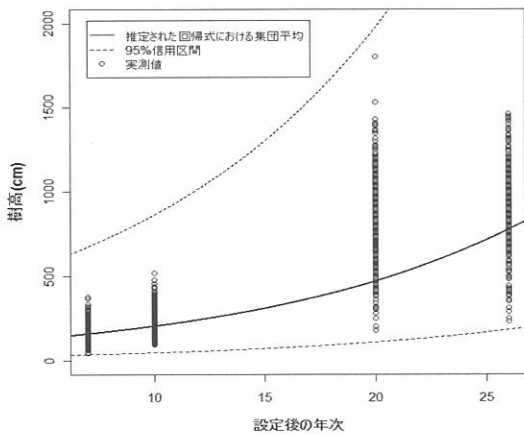


図-1.推定された回帰式と樹高測定値

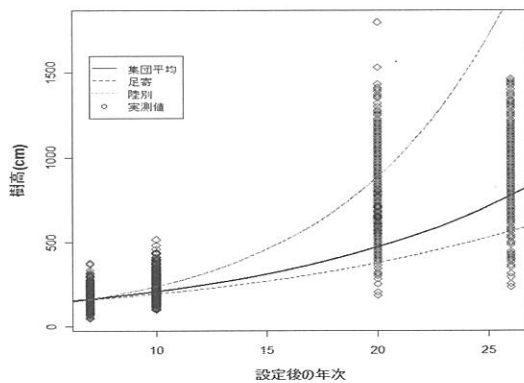


図-2. 推定された産地別の回帰式

- (2) 藤越 康祝(2009)経時データ解析の数理, 6-10, (シリーズ 多変量データの統計科学, 朝倉書店, 東京

- (3) R Development Core Team(2011) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org>.

- (3) 柳原 宏和・吉本 敦・能本 美穂(2004)林分成長分析のための一般化非線形混合効果モデル. 森林資源管理と数理モデル Vol. 3, -FORMATH TSUKUBA 2003-(鹿又秀聡・吉本敦 編), 14-46, 森林計画学会出版局, 東京.

表-1.回帰式における各要因の推定値

要因	平均	95%信用区間	
		下限	上限
産地の切片	0.236	0.084	0.482
産地の年次に対するランダム回帰係数	0.191	0.077	0.332
産地内家系の切片	0.083	0.053	0.118
年次と産地内家系との共分散	-0.002	-0.021	0.015
産地内家系の年次に対するランダム回帰係数	0.041	0.028	0.058
残差	0.088	0.084	0.093

表-2.産地の年次へのランダム係数および切片の推定値

産地	年次へのランダム係数	切片
栄浦	0.016	0.164
温根湯	0.002	-0.045
常呂	0.033	-0.128
石狩	0.030	-0.051
足寄	0.018	0.031
大樹	0.026	0.328
中川	0.014	-0.079
定山溪	0.014	0.037
東藻琴	0.003	0.015
野幌	0.018	0.010
雄武	0.007	0.057
陸別	0.023	-0.352