複数の成長パターンを持つスギ単純同齢林における炭素固定量予測

Predicting Carbon Sequestered in an Even-Aged Sugi Forest Stand through Growth Pattern Classification

柳原 宏和・吉本　敦・二宮 嘉行
Yanagihara, H., Yoshimoto, A. & Ninomiya, Y.

キーワード:　外挿での予測評価，k-平均法，情報量規準，正規多変量線形回帰モデル，成長分析，モデル選択

要約:　本研究では，複数の成長パターンが観察される林分での炭素固定量の予測手法を提示した．その方法は以下の通りである．まず，サンプル木の材積成長データに成長曲線をあてはめパラメータの推定を行い，推定されたパラメータの値を新たな観測値とする．この観測値に対してk-平均法に基づくクラスタリングを行いサンプル木の成長パターンを分類し，そのクラスター分割を考慮に入れた正規多変量線形回帰モデルをあてはめる．次に，回帰モデルのパラメータ推定値により Predictive Akaike's Information Criterion (PAIC) を計算し，PAICに基づき最適な残存木の成長パターンの分類とクラスターの個数を決定する．最後に，最適な予測モデルから残存木の成長曲線のパラメータの予測値を求め，残存木の将来的な炭素固定量の予測とその漸近$1-\alpha$信頼区間を求める．

Abstract:　In this paper, we presented a statistical procedure of estimating the clustered multivariate linear regression model for predicting the amount of carbon sequestered in a forest stand where there exist several growth patterns. The procedure is as follows: 1) By fitting a volume growth curve to the data of each sampled tree, parameters of the applied growth curve model are estimated. 2) By setting the estimated parameters as new observations, we

classify growth patterns of sampled trees by k-means method based on the new observations. 3) We construct a multivariate normal linear regression model with dummy variables for k-clusters. 4) Among a set of the estimated regression models with the different number of clusters, the best model is selected by minimizing the resultant predictive Akaike's information criterion (PAIC) for the remaining trees. 5) Finally, by using the best set of parameters of growth curves for the remaining trees, we predict the amount of carbon sequestered by remaining trees with its asymptotically $1-\alpha$ confidence interval.

Keywords: evaluation of prediction in interpolation, growth analysis, information criterion, k-means clustering, model selection, multivariate normal linear regression model

1. はじめに

2005年2月の京都議定書の発効に伴い森林の持つ炭素吸収機能が益々重要視され，森林が持つ炭素固定量を把握する必要性が増してきている．吸収される炭素固定量は森林の単位面積当たりの総材積量(m^3/ha)にある定数を掛けることにより概算することができるため，総材積の成長量を予測することが将来的に固定される炭素量の把握に必要不可欠になる．その際，仮に同一林分内に複数の成長パターンが存在すれば，それらの成長パターンを十分に考慮した予測を行う必要が出てくる．実際に成長パターンの識別はk-平均法(MacQueen 1967)に代表されるクラスタリング手法により可能であるが，分類対象となる林木はデータ収集のためにすでに伐採されている．すなわち将来的に固定される炭素量を予測するためには残存木の成長量を予測することが必要不可欠となる．そこで本研究は，複数の成長パターンが同時に存在する林分において，材積の成長量を予測する手法を開発し，残存木により固定される炭素量の予測を試みる．

予測の手順は以下の通りである．まず，データ収集のために伐採されたサンプル木に対し材積成長曲線をあてはめパラメータの推定を行う．次に，柳原・吉本(2005)，吉本ら(2005)に示されているように，推定されるパラメータの値を新たな観測値とし，それらに対しk-平均法に基づくクラスタリングを行い，サンプル木の成長パターンを分類する．次に新たな観測値を応答変数(response variable)とし，ダミー変数を用いることによりクラスター分割

を表現した正規多変量線形回帰モデルをあてはめる．残存木に対する最適な予測モデルの選択には，Predictive Akaike's Information Criterion (PAIC; Satoh 1997)を用いて残存木の成長パターンの分類と最適なクラスター数の決定を同時に行う．最後に，最適な予測モデルを用いて残存木の材積成長曲線のパラメータを推定し，残存木の炭素固定量の予測値とその漸近$1-\alpha$信頼区間を求める．

2. サンプル木のクラスター分割の決定

今，単純同齢林内のn個のサンプル林木から材積成長量のデータが得られたとし，v_{il}を第i番目の林木のt_{il}時点（樹齢）における観測値$(i=1,...,n, l=1,...,p_i)$とする．また，$\boldsymbol{v}_i=(v_{i1},...,v_{ip_i})'$を$i$番目の林木に対する成長データベクトルとする．ここで，この観測値ベクトル$\boldsymbol{v}_i$に以下のような観測時点に関する非線形な成長曲線モデルを仮定する．

[1] $$\boldsymbol{v}_i=\boldsymbol{\eta}(\boldsymbol{t}_i\,|\,\boldsymbol{\theta}_i)+\boldsymbol{\varepsilon}_i \quad (i=1,...,n)$$

ただし，$\boldsymbol{\eta}(\boldsymbol{t}_i\,|\,\boldsymbol{\theta}_i)$は各林木に対し$q$個の要素からなる未知パラメータベクトル$\boldsymbol{\theta}_i=(\theta_{i1},...,\theta_{iq})'$を持ち，時点ベクトル$\boldsymbol{t}_i=(t_{i1},...,t_{ip_i})'$の関数で定義される$p_i\times 1$平均値ベクトルである．また，$\boldsymbol{\varepsilon}_i=(\varepsilon_{i1},...,\varepsilon_{ip_i})'$は，それぞれ独立に平均$E[\boldsymbol{\varepsilon}_i]=\boldsymbol{0}_{p_i}$，分散共分散行列$\mathrm{Cov}[\boldsymbol{\varepsilon}_i]=\sigma_i^2\boldsymbol{I}_{p_i}$を持つ分布に従う誤差ベクトルとする．なお，$\boldsymbol{\eta}(\boldsymbol{t}_i\,|\,\boldsymbol{\theta}_i)$は仮定する既知関数$f(t_{ij}\,|\,\boldsymbol{\theta}_i)$により以下のように定義するものとする．

[2] $$\boldsymbol{\eta}(\boldsymbol{t}_i\,|\,\boldsymbol{\theta}_i)=\begin{pmatrix} f(t_{i1}\,|\,\boldsymbol{\theta}_i)\\ \vdots \\ f(t_{ip_i}\,|\,\boldsymbol{\theta}_i)\end{pmatrix}$$

仮に，柳原・吉本(2005)で用いたリチャーズの成長関数(Richards 1958)を用いれば，

[3] $$f(t_{il}\,|\,\boldsymbol{\theta}_i)=e^{\theta_{i1}}\left\{1-\exp\left(-e^{\theta_{i2}}t_{il}\right)\right\}^{\exp(\theta_{i3})}$$

となり，未知パラメータの個数は$q=3$となる．次に[1]式の推定量$\hat{\boldsymbol{\theta}}_i$を推定し，柳原・吉本(2005)や吉本ら(2005)と同様に，得られる推定値$\hat{\theta}_{i1},...,\hat{\theta}_{iq}$

を新たな観測値 $\boldsymbol{y}_i = (\hat{\theta}_{i1},...,\hat{\theta}_{iq})'$ とみなし，$\boldsymbol{Y} = (\boldsymbol{y}_1,...,\boldsymbol{y}_n)'$ に対して成長パターンの分類を行う．

成長パターン分類については以下の通りである．林分内には k 個の成長パターン (クラスター) が存在し，そのクラスター分割を $G = \{C_1,...,C_k\}$ とする．ただし，第 g 番目のクラスターに属する固体の数は n_g とする．なお $n_1 + \cdots + n_k = n$ である．このとき，各クラスター C_g $(g = 1,...,k)$ の重心 $\overline{\boldsymbol{y}}_g$ と共分散行列 $\boldsymbol{S}_g$ は，

$$\overline{\boldsymbol{y}}_g = \frac{1}{n_g}\sum_{i\in C_g}\boldsymbol{y}_i, \qquad \boldsymbol{S}_g = \frac{1}{n_g}\sum_{i\in C_g}(\boldsymbol{y}_i - \overline{\boldsymbol{y}}_g)(\boldsymbol{y}_i - \overline{\boldsymbol{y}}_g)' \tag{4}$$

である．MacQueen(1967)により提案された k-平均法は，以下のように分散分析における群内平方和と類似したクラスター内平方和積和行列 $\boldsymbol{W}(G) = n_1\boldsymbol{S}_1 + \cdots + n_k\boldsymbol{S}_k$ の変化量の増減により個体 i が属するクラスターを決定する．第 i 番目の個体がクラスター C_g に属しているようなクラスター分割 $G = \{C_1,...,C_k\}$ において，この個体を他のクラスター C_h $(h \neq g)$ に移動させるとする．このとき，新しいクラスター分割 $G^* = \{C_1,...,C_h^*,...,C_g^*,...,C_k\}$ におけるクラスター内平方和積和行列 $\boldsymbol{W}(G^*)$ を計算し，個体 i をクラスター C_h に移動させる前の分割 G に基づく $\boldsymbol{W}(G)$ に比べ $\boldsymbol{W}(G^*)$ が小さくなっていれば分割を更新し，そうでなければ更新しない．このようにクラスター分割の更新を逐次判定し，分割が収束するまで続ける．

上記のようにクラスター分割を更新する前のクラスター内平方和積和行列 $\boldsymbol{W}(G)$ と分割を更新した後でのクラスター内平方和積和行列 $\boldsymbol{W}(G^*)$ の差の増減により分割を更新するか否かを決定する手法が k-平均法であるが，実際 $\boldsymbol{W}(G)$ と $\boldsymbol{W}(G^*)$ は行列であるため，大きさの測り方には様々な基準がある．一般的な手法は行列のトレースにより大きさを測り，

$$\mathrm{tr}(\boldsymbol{W}(G)) > \mathrm{tr}(\boldsymbol{W}(G^*)) \tag{5}$$

であればクラスター分割を更新するというものである．しかしながら，このトレースによる更新判定は単純に個体 i とクラスター重心とのユークリッド距離の大小で行われ，同一個体内の変数同士が強い相関を持つ場合，この判定基準では相関を考慮したクラスター分割が困難となる．すなわちトレース

のみで大きさを比較する場合，単純にユークリッド距離が近いもの同士を同一のクラスターと識別するからである．そこで本論文では強い相関を考慮できるように，行列式により大きさを判断する手法を用いる．すなわち，

[6]　$|\boldsymbol{W}(G)|>|\boldsymbol{W}(G^*)|$

であればクラスター分割を更新するという手法である．

クラスター分割を更新するか否かを判断するためには，逐次更新される新しい分割でのクラスター内平方和積和行列 $\boldsymbol{W}(G^*)$ を求める必要がある．しかしながら，栁原・吉本(2005)に示されているように，それぞれの分割におけるクラスター内平方和積和行列 $\boldsymbol{W}(G)$ と $\boldsymbol{W}(G^*)$ の行列式の比，$|\boldsymbol{W}(G^*)|/|\boldsymbol{W}(G)|$ は，

[7]　$\boldsymbol{a}_g=\sqrt{\dfrac{n_g}{n_g-1}}\boldsymbol{W}(G)^{-1/2}(\boldsymbol{y}_i-\overline{\boldsymbol{y}}_g)\quad \boldsymbol{a}_h=\sqrt{\dfrac{n_h}{n_h+1}}\boldsymbol{W}(G)^{-1/2}(\boldsymbol{y}_i-\overline{\boldsymbol{y}}_h)$

とおくと，

[8]　$\dfrac{|\boldsymbol{W}(G^*)|}{|\boldsymbol{W}(G)|}=(1+\boldsymbol{a}_h{}'\boldsymbol{a}_h)(1-\boldsymbol{a}_g{}'\boldsymbol{a}_g)+(\boldsymbol{a}_g{}'\boldsymbol{a}_h)^2$

と表すことができるため，基となる分割から得られる情報のみにより分割の更新判別が可能となる．すなわち不等式 $|\boldsymbol{W}(G)|>|\boldsymbol{W}(G^*)|$ と $(1+\boldsymbol{a}_h{}'\boldsymbol{a}_h)(1-\boldsymbol{a}_g{}'\boldsymbol{a}_g)+(\boldsymbol{a}_g{}'\boldsymbol{a}_h)^2<1$ は同値となるからである．なお，$\overline{\boldsymbol{y}}_g$ と $\overline{\boldsymbol{y}}_h$ は [4] 式で与えられる，分割更新前のクラスター C_g と C_h の重心である．

上記の行列式に基づく k-平均法によるクラスタリングの手順は，以下のようになる．

<u>行列式を用いたk-平均法</u>

Step 1:　クラスターの個数 k をあらかじめ与え，乱数により各個体の分類し，初期分割を決定する．

Step 2:　第 i 番目の個体が属するクラスターを C_g とし，クラスター C_g から C_h $(h\neq g)$ に移動させる状況を考える．このとき，

[9]　$(1+\boldsymbol{a}_h{}'\boldsymbol{a}_h)(1-\boldsymbol{a}_g{}'\boldsymbol{a}_g)+(\boldsymbol{a}_g{}'\boldsymbol{a}_h)^2<1$

であれば個体 i の属するクラスターを C_g から C_h に更新する．ただし $\boldsymbol{a}_g$ と $\boldsymbol{a}_h$ は[7]式で与えられるベクトルである．

Step 3:　Step 2 の作業を全ての個体に対し，すべてのクラスターについて収

束するまで行い，最終的にそれぞれが属するクラスターを決定する．

Step 4: 初期分割を変え十分にStep 1 - Step 3を繰り返し，それぞれ得られた分割でのクラスター内平方和積和行列の行列式$|\boldsymbol{W}(G)|$が最小となるクラスター分割を最終的に最適な分割とする．

k-平均法はクラスター内平方和積和行列の行列式が最小になるようにクラスター分割Gを更新する手法であり，必ず極小値に収束するが，必ずしも最小値である保証がない．そこで，初期分割を変え十分繰り返すStep 4が必要不可欠となる．

3. 伐採木のクラスター分割と最適なクラスター数の決定法

推定量$\boldsymbol{Y}$についてk-平均法を適用しk個のクラスター分割$G=\{C_1,...,C_k\}$が得られているとする．クラスター毎に成長パターンが異なるため，クラスター毎に異なる線形モデルをあてはめる必要がある．すなわち，個体iがクラスターC_gに属するとき，$\boldsymbol{y}_i$に以下のような$r\times1$説明変数ベクトル$\boldsymbol{x}_i=(x_{i1},...,x_{ir})'$を持つ正規多変量線形回帰モデルをあてはめることになる．

[10] $$M_k:\boldsymbol{y}_i\sim i.d.\ N_q(\boldsymbol{\Xi}_g{}'\boldsymbol{x}_i,\boldsymbol{\Sigma})\quad(i=1,...,n;g=1,...,k)$$

ここで$\boldsymbol{\Xi}_g$と$\boldsymbol{\Sigma}$はそれぞれ$r\times q$と$q\times q$の平均構造と分散共分散行列を表す未知パラメータ行列であり，$\boldsymbol{\Xi}_g$の成分は以下の通りである．

[11] $$\boldsymbol{\Xi}_g=\begin{pmatrix}\xi_{11}^g & \cdots & \xi_{1q}^g\\ \vdots & \ddots & \vdots\\ \xi_{r1}^g & \cdots & \xi_{rq}^g\end{pmatrix}$$

次に$\boldsymbol{\Xi}'=(\boldsymbol{\Xi}_1',...,\boldsymbol{\Xi}_k')$とし，$\boldsymbol{\Xi}$を用いてクラスター毎に異なるモデルに対しダミー変数を用いて一つのモデルで表す．今，個体iがクラスターC_gに属せば1，属さないのであれば0であるようなダミー変数d_{ig}を考え，そのベクトルを$\boldsymbol{d}_i=(d_{i1},...,d_{ik})'$とする．このとき，個体が$g$番目のクラスターに属するため，$\boldsymbol{d}_i$は$g$番目の成分が1で残りが0となる$k\times1$ベクトルである．このとき，ダミー変数を加えて新たに$kr\times1$説明変数ベクトル

$\boldsymbol{x}_{d,i} = \boldsymbol{d}_i \otimes \boldsymbol{x}_i$ を考える．ただし“⊗”はクロネッカー積を表す．なお，$n \times m$ 行列 $\boldsymbol{A} = [a_{ij}]$ と行列 $\boldsymbol{B}$ のクロネッカー積は，

[12] $$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{pmatrix} a_{11}\boldsymbol{B} & \cdots & a_{1m}\boldsymbol{B} \\ \vdots & \ddots & \vdots \\ a_{n1}\boldsymbol{B} & \cdots & a_{nm}\boldsymbol{B} \end{pmatrix}$$

である(参照 Magnus and Neudecker 1999, p.27-28)．従って $\boldsymbol{x}_{d,i}$ は，

[13] $$\boldsymbol{x}_{d,i} = \boldsymbol{d}_i \otimes \boldsymbol{x}_i = \begin{pmatrix} \boldsymbol{0}_{(g-1)r} \\ \boldsymbol{x}_i \\ \boldsymbol{0}_{(k-g)r} \end{pmatrix}$$

となる．この説明変数ベクトル $\boldsymbol{x}_{d,i}$ を用いると，

[14] $$\boldsymbol{\Xi}'\boldsymbol{x}_{d,i} = \begin{pmatrix} \boldsymbol{\Xi}_1' & \cdots & \boldsymbol{\Xi}_g' & \cdots & \boldsymbol{\Xi}_k' \end{pmatrix} \begin{pmatrix} \boldsymbol{0}_r \\ \vdots \\ \boldsymbol{x}_i \\ \vdots \\ \boldsymbol{0}_r \end{pmatrix} = \boldsymbol{\Xi}_g'\boldsymbol{x}_i$$

となることから，[10] 式は

[15] $$M_k : \boldsymbol{y}_i \sim i.d.\ N_q(\boldsymbol{\Xi}'\boldsymbol{x}_{d,i}, \boldsymbol{\Sigma}) \quad (i = 1,\ldots,n)$$

と表すことができる．このとき，モデル M_k における $\boldsymbol{y}_i$ の確率密度関数は，

[16] $$L(\boldsymbol{y}_i \mid \boldsymbol{x}_{d,i}, \boldsymbol{\Xi}, \boldsymbol{\Sigma}) = \left(\frac{1}{2\pi}\right)^{q/2} |\boldsymbol{\Sigma}|^{-1/2} \exp\left\{(\boldsymbol{y}_i - \boldsymbol{\Xi}'\boldsymbol{x}_{d,i})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_i - \boldsymbol{\Xi}'\boldsymbol{x}_{d,i})\right\}$$

である(参照 塩谷 1990, p.7-12)．今，$n \times kr$ 行列 $\boldsymbol{X}_d$ を $\boldsymbol{X}_d = (\boldsymbol{x}_{d,1},\ldots,\boldsymbol{x}_{d,n})'$ と置くと，モデル M_k の下での $\boldsymbol{\Xi}$ と $\boldsymbol{\Sigma}$ の最尤推定量は，

[17] $$\hat{\boldsymbol{\Xi}}_Y = (\boldsymbol{X}_d'\boldsymbol{X}_d)^{-1}\boldsymbol{X}_d'\boldsymbol{Y}, \quad \hat{\boldsymbol{\Sigma}}_Y = \frac{1}{n}\boldsymbol{Y}'\left[\boldsymbol{I}_n - \boldsymbol{X}_d(\boldsymbol{X}_d'\boldsymbol{X}_d)^{-1}\boldsymbol{X}_d'\right]\boldsymbol{Y}$$

となる．

実際に林木の成長データを用いる際，上記で求められるモデル M_k は伐採されたサンプル木，つまり実測値に基づくモデルであり，残存木に関するモデルではない．すなわち残存木のクラスター分割を評価するためには，M_k のモデル評価ではなく残存木に関するモデル評価を行わなければならない．

そのためには，残存木の成長曲線のパラメータに関する予測モデルを次のように考える必要がある．

今，残存木の本数を m として，$\boldsymbol{\zeta}_j$ と $\boldsymbol{\delta}_j$ $(j=1,...,m)$ を残存木の材積成長曲線のパラメータベクトルとクラスター分割を表現する観測されていないダミー変数ベクトルとする．また観測される残存木の説明変数ベクトルを $\boldsymbol{w}_j$ とし，$\boldsymbol{w}_{\delta,j}=\boldsymbol{\delta}_j\otimes\boldsymbol{w}_j$ とする．このとき，残存木もサンプル木から得られる実測値モデル M_k と同様のプロセスに従うと仮定し，残存木に関する予測モデルを

[18] $$\mathcal{M}_k:\boldsymbol{\zeta}_j\sim i.d.N_q(\boldsymbol{\Xi}'\boldsymbol{w}_{\delta,j},\boldsymbol{\Sigma})\quad(j=1,...,m)$$

とする．なお，予測モデル $\mathcal{M}_k$ において，$\boldsymbol{\Xi}$ と $\boldsymbol{\Sigma}$ は実測値モデル M_k と共通である．

最適な実測値モデル M_k の選択には予測カルバック－ライブラーの距離(Kullback and Leibler 1951)に基づくリスクを用いることが一般的である．そのリスクは，$\boldsymbol{Y}$ と独立で同一な分布に従う確率変数 $\boldsymbol{Y}_{\mathrm{F}}=(\boldsymbol{y}_{\mathrm{F},1},...,\boldsymbol{y}_{\mathrm{F},n})'$ を用いて，

[19] $$R_k(\boldsymbol{X}_d\mid\boldsymbol{X}_d)=-2\sum_{i=1}^{n}E_{\boldsymbol{Y}}^{*}E_{\boldsymbol{Y}_{\mathrm{F}}}^{*}\left[\log L(\boldsymbol{y}_{\mathrm{F},i}\mid\boldsymbol{x}_{d,i},\hat{\boldsymbol{\Xi}}_{\boldsymbol{Y}},\hat{\boldsymbol{\Sigma}}_{\boldsymbol{Y}})\right]$$

と定義される．ただし E^{*} は真のモデルの下での期待値である．実測値モデルにおける予測リスクに関しては，吉本ら(2005)に詳しい．上記のリスクでは，$\boldsymbol{X}_d$ という説明変数が与えられたとき，その $\boldsymbol{X}_d$ の下で予測を行ったという状況でモデルを評価する．すなわちモデルの内挿での予測に基づく評価である．しかしながら，ここで評価するモデルは $\mathcal{M}_k$ であり，この場合，$\boldsymbol{\mathcal{Z}}=(\boldsymbol{\zeta}_1,...,\boldsymbol{\zeta}_m)'$ は観測されないため，$\boldsymbol{\Xi}$ と $\boldsymbol{\Sigma}$ の推定は $\boldsymbol{Y}$ から行わなくてはならない．その結果，予測に基づくリスクは，

[20] $$R_k(\boldsymbol{W}_{\delta}\mid\boldsymbol{X}_d)=-2\sum_{j=1}^{m}E_{\boldsymbol{Y}}^{*}E_{\boldsymbol{\mathcal{Z}}}^{*}\left[\log L(\boldsymbol{\zeta}_j\mid\boldsymbol{w}_{\delta,j},\hat{\boldsymbol{\Xi}}_{\boldsymbol{Y}},\hat{\boldsymbol{\Sigma}}_{\boldsymbol{Y}})\right]$$

となる．ただし，$\boldsymbol{W}_{\delta}=(\boldsymbol{w}_{\delta,1},...,\boldsymbol{w}_{\delta,m})'$ である．このリスクでは，$\boldsymbol{X}_d$ と異なる $\boldsymbol{W}_{\delta}$ の下で予測を行った場合のモデルを評価することになり，モデルの外挿での予測に基づくモデル評価である．仮に $\boldsymbol{W}_{\delta}=\boldsymbol{X}_d$ となれば，[19] 式と

[20] 式は一致する．以下ではリスク $R_k(\boldsymbol{W}_\delta \mid \boldsymbol{X}_d)$ の最小化により最適な予測モデル $\mathcal{M}_k$ の選択を行う．

上記の [20] 式のリスクの推定量は Satoh(1997)により提案されたPredictive Akaike's Information Criterion (PAIC)である．予測モデル $\mathcal{M}_k$ に関する PAIC は以下のように定義される．

[21] $$\mathrm{PAIC}_k = m\log|\hat{\boldsymbol{\Sigma}}_Y| + mq\log 2\pi + \frac{n(m+\phi_k)q}{n-kr-q-1}$$

ただし，

[22] $$\phi_k = \mathrm{tr}\left\{\boldsymbol{W}_\delta'\boldsymbol{W}_\delta(\boldsymbol{X}_d'\boldsymbol{X}_d)^{-1}\right\} = \sum_{j=1}^{m}(\boldsymbol{\delta}_j'\otimes\boldsymbol{w}_j')(\boldsymbol{X}_d'\boldsymbol{X}_d)^{-1}(\boldsymbol{\delta}_j\otimes\boldsymbol{w}_j)$$

ここで，クラスター数 k を固定して考えると，$\hat{\boldsymbol{\Xi}}_Y$ と $\hat{\boldsymbol{\Sigma}}_Y$ は一定となり，PAIC の最小化は ϕ_k の最小化に等しくなる．このとき ϕ_k の値は残存木のクラスター分割 $\boldsymbol{\delta}_1,\ldots,\boldsymbol{\delta}_m$ のみによって決定されるため，ϕ_k を最小にする $\boldsymbol{\delta}_1,\ldots,\boldsymbol{\delta}_m$ はそのまま PAIC を最小にするクラスター分割となる．よって，k を固定した下での残存木の最適なクラスター分割は，ϕ_k を最小にする $\boldsymbol{\delta}_1,\ldots,\boldsymbol{\delta}_m$ になる．個々の $(\boldsymbol{\delta}_j'\otimes\boldsymbol{w}_j')(\boldsymbol{X}_d'\boldsymbol{X}_d)^{-1}(\boldsymbol{\delta}_j\otimes\boldsymbol{w}_j)$ は正定値行列の二次形式であり，その値は常に正になることから，ϕ_k を最小にするためには $(\boldsymbol{\delta}_j'\otimes\boldsymbol{w}_j')(\boldsymbol{X}_d'\boldsymbol{X}_d)^{-1}(\boldsymbol{\delta}_j\otimes\boldsymbol{w}_j)$ を最も小さくするように $\boldsymbol{\delta}_j$ を決定すればよいことになる．すなわち，最適なクラスター分割は，g 番目の成分が 1 で残りが 0 であるような $k\times 1$ ベクトル $\boldsymbol{e}_g$ を用いると，

[23] $$\hat{\boldsymbol{\delta}}_j = \arg\min_{\boldsymbol{e}=\boldsymbol{e}_1,\ldots,\boldsymbol{e}_k}(\boldsymbol{e}'\otimes\boldsymbol{w}_j')(\boldsymbol{X}_d'\boldsymbol{X}_d)^{-1}(\boldsymbol{e}\otimes\boldsymbol{w}_j)\ (j=1,\ldots,m)$$

となる．更に，残存木に対するクラスターの個数を変えてそれぞれのクラスターの個数でのPAICを比較することにより最適なクラスター数も決定する．つまり，それぞれの k の下でPAICが最小となるクラスター分割 $\hat{\boldsymbol{\delta}}_1,\ldots,\hat{\boldsymbol{\delta}}_m$ を決定し，その分割の下でのPAICを求め，PAICを最小とする k を最適なクラスター数とみなす．以下に残存木に対するクラスター分割と最適なクラスター数を決定する手順を示す．

残存木のクラスター分割と最適なクラスターの個数の決定法

Step 1:　最大のクラスター数 K を決め，$k=1,\ldots,K$ と繰り返す．

Step 2:　伐採された個々のサンプル木の材積成長データに成長曲線をあては

め，その推定された係数を新しい観測値 $\boldsymbol{Y}$ とする．

Step 3: クラスターの個数 k に対し，$\boldsymbol{Y}$ にクラスター内平方和積和行列の行列式の最小化による k-平均法を適用し，サンプル木の最適なクラスター分割 $\boldsymbol{d}_1,...,\boldsymbol{d}_n$ を決定する．

Step 4: 説明変数行列を $\boldsymbol{X}_d=(\boldsymbol{d}_1\otimes\boldsymbol{x}_1,...,\boldsymbol{d}_n\otimes\boldsymbol{x}_n)'$ とし，$\hat{\boldsymbol{\Xi}}_{\boldsymbol{Y}}$ と $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{Y}}$ を求め，ϕ_k の最小化により残存木のクラスター分割 $\hat{\boldsymbol{\delta}}_1,...,\hat{\boldsymbol{\delta}}_m$ を決定する．これらを用いて PAIC を計算する．

Step 5: Step 3 と 4 を繰り返し，以下のようにPAICを最小にする k を最適なクラスター数 k_{opt} とする．

[24] $$k_{\text{opt}}=\arg\min_{k=1,...,K}\text{PAIC}_k$$

4. 炭素固定量の予測

ここでは，n 本の林木の材積成長データが得られている面積 S (ha) を持つ試験林に，m 本の林木が伐採されずに残っているとし，この残存木の炭素固定量を予測する．スギ林分の炭素固定量は以下のような単位あたりの総材積により求める(参照 松本 2001).

[25] $$\text{炭素固定量 }(\text{C}t/ha)=\text{総幹材積 }(m^3)\times 0.38(\text{比重})\times 0.44(\text{スギ炭素割合})\times\frac{100}{61}\times\frac{1}{\text{林分面積 }(ha)}$$

すなわち，林分内の残存木の総材積量を予測できれば，炭素固定量も同時に予測することができる．

まず，残存木での総材積はそれぞれの林木の材積成長曲線の総和によって予測する．このとき，残存木の材積成長曲線のパラメータを前章と同様に $\boldsymbol{\zeta}_j$ $(j=1,...,m)$ と置くと，時点 t における炭素固定量の予測値は，

[26] $$\mathcal{G}_{\text{cs}}(t)=\frac{16.72}{61S}\sum_{j=1}^{m}f(t\,|\,\boldsymbol{\zeta}_j)\quad(\text{C}t/ha)$$

となる．しかしながら，$\boldsymbol{\zeta}_j$ は観測されない変数であり，その値を実際に用いることはできないため，予測モデルにより $\boldsymbol{\zeta}_j$ を予測し，その予測値を用いて [26] 式の推定量を求める．今，$\hat{\boldsymbol{\Xi}}_{\boldsymbol{Y}}$ と $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{Y}}$ をPAICの最小化により選ばれた最適なモデルの $\boldsymbol{\Xi}$ と $\boldsymbol{\Sigma}$ の推定量，$\hat{\boldsymbol{\delta}}_1,...,\hat{\boldsymbol{\delta}}_m$ を残存木の最適なモデルでの

最適なクラスター分割とする．このとき $\boldsymbol{w}_{\delta\cdot j} = \hat{\boldsymbol{\delta}}_j \otimes \boldsymbol{w}_j$ とすると，ζ_j の予測値は $\hat{\zeta}_j = \hat{\boldsymbol{\Xi}}_Y{}' \boldsymbol{w}_{\delta\cdot j}$ となる．$\mathcal{G}_{cs}(t)$ の推定量は，[26] 式で観測されない ζ_j を予測値 $\hat{\zeta}_j$ で置き換えることにより求めることができ，その推定量は以下のようになる．

[27] $$\hat{\mathcal{G}}_{cs}(t) = \frac{16.72}{61S} \sum_{j=1}^{m} f(t \mid \hat{\boldsymbol{\Xi}}_Y{}' \boldsymbol{w}_{\delta\cdot j}) \quad (\mathrm{C}t/ha)$$

この推定量 $\hat{\mathcal{G}}_{cs}(t)$ の値は実測値 $\boldsymbol{Y}$ によって変動するため，推定値の変動を考慮した推定も行うことが望ましい．その場合，信頼係数 $1-\alpha$ の区間推定を行えばよいが，成長曲線モデルが非線形関数であるため，信頼係数が正確に $1-\alpha$ となる信頼区間を構成することはできない．そのため，信頼係数は漸近的なものによって代用することにする．今，z_α を標準正規分布の上側 $100\times\alpha$ % 点とし，偏微分に関する $q\times 1$ ベクトル $\boldsymbol{h}_j(t \mid \boldsymbol{\Xi}_0)$ と $krq\times 1$ ベクトル $\boldsymbol{g}_j(t \mid \boldsymbol{\Xi}_0)$ を以下のように定義する．

[28]
$$\boldsymbol{h}_j(t \mid \boldsymbol{\Xi}_0) = \frac{\partial}{\partial \boldsymbol{\zeta}} f(t \mid \boldsymbol{\zeta}) \Big|_{\boldsymbol{\zeta} = \boldsymbol{\Xi}_0{}' \boldsymbol{w}_{\delta\cdot j}}$$
$$\boldsymbol{g}_j(t \mid \boldsymbol{\Xi}_0) = \frac{\partial}{\partial \mathrm{vec}(\boldsymbol{\Xi})} f(t \mid \boldsymbol{\Xi}' \boldsymbol{w}_{\delta\cdot j}) \Big|_{\boldsymbol{\Xi} = \boldsymbol{\Xi}_0} = (\boldsymbol{I}_q \otimes \boldsymbol{w}_{\delta\cdot j}) \boldsymbol{h}_j(t \mid \boldsymbol{\Xi}_0)$$

このとき，

[29] $$\bar{\boldsymbol{g}}(t \mid \boldsymbol{\Xi}_0) = \frac{1}{m} \sum_{j=1}^{m} \boldsymbol{g}_j(t \mid \boldsymbol{\Xi}_0) \quad \bar{\boldsymbol{H}}(t \mid \boldsymbol{\Xi}_0) = \frac{1}{m} \sum_{j=1}^{m} \boldsymbol{h}_j(t \mid \boldsymbol{\Xi}_0) \boldsymbol{h}_j(t \mid \boldsymbol{\Xi}_0)'$$

とし，

[30] $$\psi^2(t \mid \boldsymbol{\Xi}_0, \boldsymbol{\Sigma}_0) = \mathrm{tr}\left[\bar{\boldsymbol{H}}(t \mid \boldsymbol{\Xi}_0) \boldsymbol{\Sigma}_0\right] + m \bar{\boldsymbol{g}}(t \mid \boldsymbol{\Xi}_0)' \left\{\boldsymbol{\Sigma}_0 \otimes (\boldsymbol{X}_d{}' \boldsymbol{X}_d)^{-1}\right\} \bar{\boldsymbol{g}}(t \mid \boldsymbol{\Xi}_0)$$

とする．ただし，$\mathrm{vec}(\boldsymbol{A})$ は行列 $\boldsymbol{A}$ の列ベクトルを縦に並べてできるベクトルを表し，$n\times m$ 行列 $\boldsymbol{A} = [a_{ij}]$ であれば，

[31] $$\mathrm{vec}(\boldsymbol{A}) = (a_{11}, a_{21}, \ldots, a_{n1}, \ldots, a_{1m}, a_{2m}, \ldots, a_{nm})'$$

となる(参照 Magnus and Neudecker 1999, p. 30-31)．このとき $\mathcal{G}_{cs}(t)$ の漸近 $1-\alpha$ 信頼区間を以下のように定義する．

[32] $$\hat{\mathcal{G}}_{cs}^{-}(t) \le \mathcal{G}_{cs}(t) \le \hat{\mathcal{G}}_{cs}^{+}(t)$$

ただし，

[33]
$$\hat{\mathcal{G}}_{cs}^{-}(t) = \hat{\mathcal{G}}_{cs}(t) - \left(\frac{16.77}{61S}\right)\sqrt{m}z_{\alpha/2}\psi(t \mid \hat{\boldsymbol{\Xi}}_{\boldsymbol{Y}}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{Y}}) \ (\mathrm{C}t/ha)$$
$$\hat{\mathcal{G}}_{cs}^{+}(t) = \hat{\mathcal{G}}_{cs}(t) + \left(\frac{16.77}{61S}\right)\sqrt{m}z_{\alpha/2}\psi(t \mid \hat{\boldsymbol{\Xi}}_{\boldsymbol{Y}}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{Y}}) \ (\mathrm{C}t/ha)$$

であり，$\psi(t \mid \boldsymbol{\Xi}_0, \boldsymbol{\Sigma}_0) = \sqrt{\psi^2(t \mid \boldsymbol{\Xi}_0, \boldsymbol{\Sigma}_0)}$ である．このとき以下のような定理が成り立つ．なお，定理の証明については Appendix に記す．

定理: $\sqrt{n/m} = O(1)$ を仮定する．このとき，

[34] $$\mathrm{P}\left(\hat{\mathcal{G}}_{cs}^{-}(t) \le \mathcal{G}_{cs}(t) \le \hat{\mathcal{G}}_{cs}^{+}(t)\right) \to 1-\alpha \ (n \to \infty)$$

が成り立つ．

5. 実データへの適用

本章では，前章で導入した手法を用いて，炭素固定量の予測を行った．今回解析に使用したデータは，柳原・吉本(2005)で用いた福岡県八女郡星野村における23年生の無間伐林より抽出した30 $(=n)$ 本のサンプル木から得た成長データである．残存木の本数は106 $(=m)$ 本であった．成長曲線としては，[3] 式でのリチャーズの成長関数を用い，その推定値を $\boldsymbol{y}_i = (\hat{\theta}_{i1}, \hat{\theta}_{i2}, \hat{\theta}_{i3})'$ $(i = 1,...,30)$ として解析に用いた．試験林の形や立木位置，また材積成長データと成長曲線のパラメータの推定値等は，柳原・吉本(2005) に詳しい．本解析において使用した説明変数は胸高直径DBHである．従って説明変数ベクトルはそれぞれ $\boldsymbol{x}_i = (1, \mathrm{DBH}_i)'$ $(i = 1,...,30)$，$\boldsymbol{w}_j = (1, \mathrm{DBH}_j)'$ $(j = 1,...,106)$ となり，$r = 2$ となる．図1に試験林での DBH

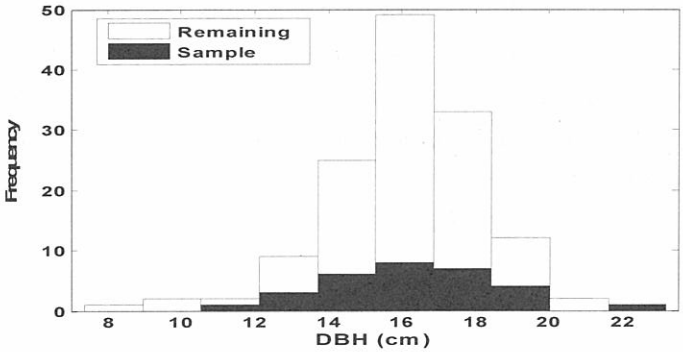

図1. サンプル木と残存木の DBH のヒストグラム

のヒストグラムを記す．それぞれ黒く塗りつぶれた棒がサンプル木のDBH，白抜きの棒が残存木のDBHを表している．この図から分布の形状は異なるものの，レンジに関してはサンプル木と伐採木でさほど変らないことがわかる．

まずサンプル木に対して，最大クラスター数 $K=5$ として行列式に基づく k-平均法を用いてクラスタリングを行った．図2.1，2.2，2.3，2.4は $k=2,...,5$ でのクラスター分割の結果である．図中の n_g $(g=1,...,k)$ はクラスター C_g に属していると判断されたサンプル木の本数を表している．なお $n_1+\cdots+n_k=n$ である． これらの図から，どの分割においても共通の 8 本の林木がクラスター 1 に分類されていることがわかる．そのため，これらの 8 本の林木は他の22本とは明らかに異なる成長パターンを持つと推測される．その他22本のクラスター分割は k の個数によって異なるが，クラスターの個数を増やせば，分割されたクラスターに属する標本の数はほぼ同じになることがわかる．

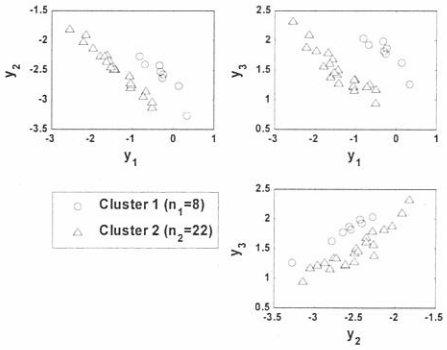

図2.1. サンプル木でのクラスタリング結果 (k=2)

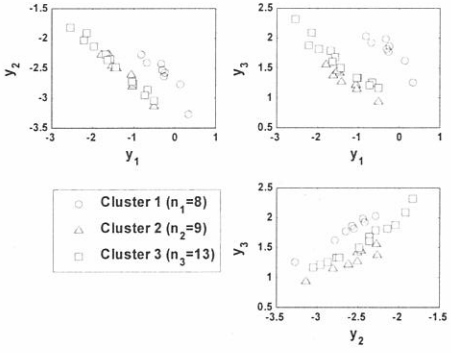

図2.2. サンプル木でのクラスタリング結果 (k=3)

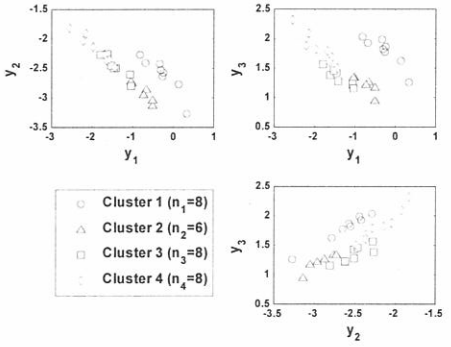

図2.3. サンプル木でのクラスタリング結果 (k=4)

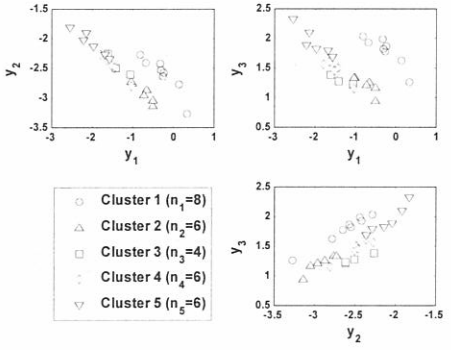

図2.4. サンプル木でのクラスタリング結果 (k=5)

次に，サンプル木のクラスタリング結果に基づき正規多変量線形モデルをあてはめ，予測モデル $\{\mathcal{M}_1,\ldots,\mathcal{M}_5\}$ に対するPAICを比較することにより最適なクラスター数と予測モデルを決定した．それぞれのモデルに対するPAICは図3のようになり，3 個のクラスターを持つモデルが最適なモデルとして選ばれた．大まかに見ると，観測値 $\boldsymbol{Y}$ のクラスターは 2 個のように観察されるが，y_1-y_2 や y_1-y_3 での散布図では不明であるものの，y_2-y_3 での散布図を見ると，3 個目のクラスターがあるように観察できる．PAICの値から，この 3 個目のクラスターもモデル化した方が良いと判断したことになる．

最適なモデルからの残存木の材積成長曲線のパラメータの予測結果を図4に示す．図中の m_g $(g=1,\ldots,k)$ はクラスター C_g に属していると判断された残存木の本数を表している．なお $m_1+\cdots+m_k=m$ である．サンプル木の場合と比べると，3 番目のクラスターに属するデータが長く伸びている．これは，サンプル木では観測されなかった非常に小さいDBHが残存木にあるからである．それらを除けばサンプル木でのクラスターとほぼ同様のクラスター分割が得られる．

残存木のクラスター分類とDBHの大きさの関係を図5に示す．この図から，大まかにDBHが比較的大きい林木は 1 番目のクラスターに，平均的な林木は 2 番目のクラスターに，そして比較的小さい林木は 3 番目のクラスターに分類されていることがわかる．

最後に，ここで得られた予測モデルを用いて計算された炭素固定量の予測結果を図6に示す．図6は 1 から 5 個のクラスターを持つモデルにおいて計算された $\hat{\mathcal{G}}_{\infty}(t)$ に対し 5 年おきに求めたグラフである．また 図7は最適なモデル $k=3$ での残存木の炭素固定量の漸近0.95信頼区間を表している．図6から，クラスターの数が増加するに伴い，炭素固定量の予測値が増えているのがわかる．その結果，もともと成長パターンが異なる林木に対して，パターンの違いを考慮しない場合，推定値の当てはまりが悪くなり，予測値を過小評価する傾向があると予想できる．逆にクラスターの個数を多めに設定すれ

ば，クラスター分類に関して誤判別を起こす危険性が高まり，予測値が増えたとしてもその値の信頼性は低くなる可能性がでてくる．

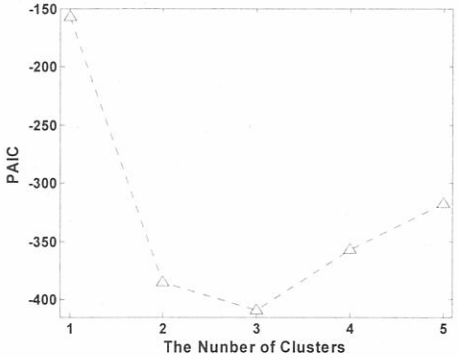

図3. クラスターの個数とPAIC

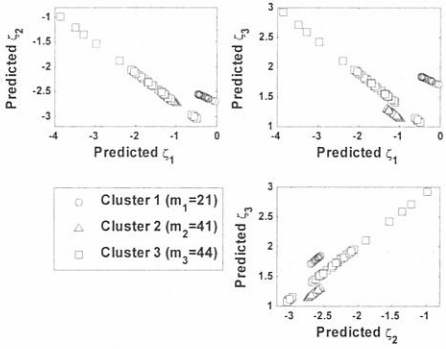

図4. 最適な予測モデルでの残存木の成長パラメータの予測値

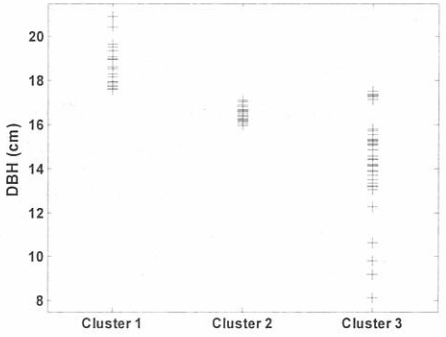

図5. 最適モデルでのクラスター分類とDBHの関係

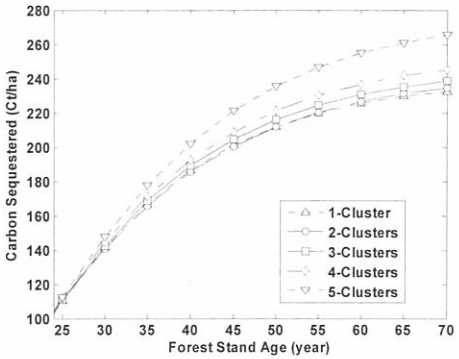

図6. それぞれのクラスターでの炭素固定量の予測量

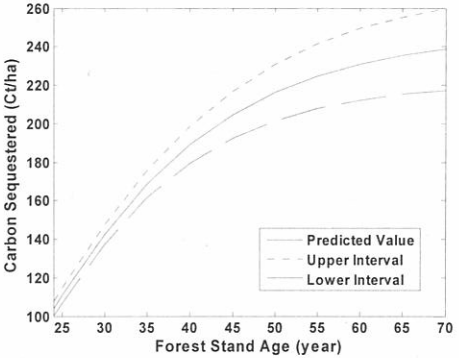

図7. 最適なモデルでの炭素固定量の予測量

6. おわりに

本論文では，同一林分の内に複数の成長パターンが存在するとき，サンプル木から残存木の成長パターンを分類し，その分割に基づいて炭素固定量を予測する手法を提示した．ここで用いた方法では，実際に残存木の成長データが得られないため，外挿での予測に基づくリスクの推定量であるPAICの最小化により，残存木のクラスター分割と最適なクラスター数，及び最適な予測モデルを決定した．ここで提示された手法と実データを用いて，残存木の成長曲線のパラメータを予測し，その値を用いて将来的に固定される炭素量の予測及びその漸近0.95信頼区間の推定を行った．

近年，一般化非線形混合効果モデル(Generalized Non-linear Mixed-effects Model; Vonesh and Carter 1992)等の多変量モデルを用いた成長データの開発が行われている．また，その種の多変量モデルを用いた炭素固定量の予測も行われている(Yanagihara and Yoshimoto 2005)．しかしながら，それらのモデルで残存木の成長パターンの分類を行える有効な手法は開発されていない．また混合効果モデルはモデルの特定のために莫大な計算量を必要とするため，候補のモデルが多い場合，計算量的にもその使用は困難である．そのような問題点に対し，今回提示した手法は簡便でかつ有効な手法であると考える．

実データを用いた分析では，説明変数としてDBHのみを用いたが，成長に影響を与える因子は他にも考えられる．吉本ら(2005)は，線形回帰モデルにおいて，周辺のDBHの平均，特に，半径 5 m以内の周りの林木のDBHの重み付き平均もDBHと同様に最適な説明変数として選択している．仮にそれらの変数を用いて内挿での予測精度が上がれば，外挿での予測精度もあがる．そのため，より精度の高い予測ができるような説明変数の組み合わせを探求し，炭素固定量の予測を行うことも今後の課題である．

謝辞

本研究は，文部科学省科学研究費（基盤研究（B)(2):No.15330048）及び地球環境研究総合推進費S-4を受けて行われたものである．

引用文献

Kullback, S. and Leibler, R. A. 1951. On information and sufficiency, *Annal of Mathematical Statistics* 22:79-86

MacQueen, J. B. 1967. Some methods for classification and analysis of multivariate observations, pp.281-298, Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability (Neyman, J. ed.), 1, Berkeley,

Magnus, J. R. and Neudecker, H. 1999. Matrix differential calculus with applications in statistics and econometrics (Revised ed.), John Wiley & Sons, New York

松本光朗. 2001. 日本の森林による炭素蓄積量と炭素吸収量, *森林科学* 33:30-36

Richards, F. J. 1958. A flexible growth function to empirical use, *Journal of Experimental Botany* 10:290-300

Satoh, K. 1997. AIC-type model selection criterion for multivariate linear regression with a future experiment, *Journal of the Japan Statistical Society* 27:135-140

Seber, G. A. F. and Wild, C. J. 1989. Nonlinear Regression, John Wiley & Sons, New York

塩谷實. 1990. 多変量解析概論, 朝倉書店, 東京

Vonesh, E. F. and Carter, R. L. 1992. Mixed-effects nonlinear regression for unbalanced repeated measures, *Biometrics* 48:1-17

栁原宏和・吉本敦. 2005. 単純同齢林における林木成長パターンのクラスタリング, pp.49–69, 森林資源管理と数理モデル Vol. 4 (近藤洋史・吉本敦・松村直人 編集), 森林計画学会出版局, 東京

Yanagihara, H. and Yoshimoto, A. 2005. Statistical procedure for assessing the amount of carbon sequestrated by sugi (*Cryptomeria japonica*) plantation, pp.125–140, Multipurpose Inventory for the Aged Artificial Forest (Nobori, Y., Takahashi, N. & Yoshimoto, A. eds.), Japan Society of Forest Planning Press, Utsunomiya,

吉本敦・栁原宏和・二宮嘉行. 2005. 多変量線形モデルによる林分成長要因探索のための変数選択, *日林誌* 87:504-512

Appendix

$\boldsymbol{y}_1,...,\boldsymbol{y}_n$ と $\boldsymbol{\zeta}_1,...,\boldsymbol{\zeta}_m$ は独立に多変量正規分布に従う以下のような真のモデルから生成される確率変数とする．

[A1]
$$M_*: \boldsymbol{y}_i \sim i.d.\ N_q(\boldsymbol{\Xi}_*{}'\boldsymbol{x}_{d,i}, \boldsymbol{\Sigma}_*)\ (i=1,...,n)$$
$$\mathcal{M}_*: \boldsymbol{\zeta}_j \sim i.d.\ N_q(\boldsymbol{\Xi}_*{}'\boldsymbol{w}_{\delta,j}, \boldsymbol{\Sigma}_*)\ (j=1,...,m)$$

ここで，$\boldsymbol{\beta}_j = \boldsymbol{\zeta}_j - \boldsymbol{\Xi}_*{}'\boldsymbol{w}_{\delta,j}$ と置くと，$\boldsymbol{\beta}_1,...,\boldsymbol{\beta}_m \sim i.i.d.\ N_q(\boldsymbol{0}_q, \boldsymbol{\Sigma}_*)$ である．このとき，非線形関数 $f(t\,|\,\boldsymbol{\zeta}_j)$ を，テーラー展開を用いた一次近似(参照 Seber and Wild 1989, p.23-25)により近似すると，

[A2]
$$f(t\,|\,\boldsymbol{\zeta}_j) \approx f(t\,|\,\boldsymbol{\Xi}_*{}'\boldsymbol{w}_{\delta,j}) + \boldsymbol{h}_j(t\,|\,\boldsymbol{\Xi}_*)'\boldsymbol{\beta}_j$$

となる．また，$krq\times 1$ ベクトル $\boldsymbol{u}$ を

[A3]
$$\boldsymbol{u} = \sqrt{n}\left\{\mathrm{vec}(\hat{\boldsymbol{\Xi}}_{\boldsymbol{Y}}) - \mathrm{vec}(\boldsymbol{\Xi}_*)\right\}$$

とすると，$\boldsymbol{Y}$ の正規性の仮定より，$\boldsymbol{u}$ は以下のような正規分布に従う．

[A4]
$$\boldsymbol{u} \sim N_{krq}(\boldsymbol{0}_{krq}, \boldsymbol{\Sigma}_* \otimes n(\boldsymbol{X}_d{}'\boldsymbol{X}_d)^{-1})$$

この変数を用いると，テーラー展開により，

[A5]
$$f(t\,|\,\hat{\boldsymbol{\Xi}}_{\boldsymbol{Y}}{}'\boldsymbol{w}_{\delta,j}) = f(t\,|\,\boldsymbol{\Xi}_*{}'\boldsymbol{w}_{\delta,j}) + \frac{1}{\sqrt{n}}\boldsymbol{g}_j(t\,|\,\boldsymbol{\Xi}_*)'\boldsymbol{u} + O_p(n^{-1})$$

となる．ここで，$\sqrt{n/m} = O(1)$ であれば，

[A6]
$$\frac{\mathcal{G}_{cs}(t) - \hat{\mathcal{G}}_{cs}(t)}{a_S\sqrt{m}} = \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\boldsymbol{h}_j(t\,|\,\boldsymbol{\Xi}_*)'\boldsymbol{\beta}_j - \sqrt{\frac{m}{n}}\bar{\boldsymbol{g}}(t\,|\,\boldsymbol{\Xi}_*)'\boldsymbol{u} + O_p(n^{-1/2})$$

と展開できる．ただし $a_S = 16.72/61S$ である．ここで $\boldsymbol{\beta}_1,...,\boldsymbol{\beta}_m$ と $\boldsymbol{u}$ は互いに独立に正規分布に従う確率変数なので，[A6] 式の右辺は正規分布に収束する．よって，

[A7]
$$\frac{\mathcal{G}_{cs}(t) - \hat{\mathcal{G}}_{cs}(t)}{a_S\sqrt{m}} \xrightarrow{\mathcal{D}} N(0, \psi^2(t\,|\,\boldsymbol{\Xi}_*, \boldsymbol{\Sigma}_*))\ \left(n\to\infty, \sqrt{n/m} = O(1)\right)$$

を満たす．このとき，$\hat{\boldsymbol{\Xi}}_{\boldsymbol{Y}} \to \boldsymbol{\Xi}_*$，$\hat{\boldsymbol{\Sigma}}_{\boldsymbol{Y}} \to \boldsymbol{\Sigma}_*$ $(n\to\infty)$ であるため，$\psi^2(t\,|\,\hat{\boldsymbol{\Xi}}_{\boldsymbol{Y}}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{Y}}) \to \psi^2(t\,|\,\boldsymbol{\Xi}_*, \boldsymbol{\Sigma}_*)$ $(n\to\infty)$ となることがわかる．よって，以下の式が成り立つ．

[A8]
$$\frac{\mathcal{G}_{cs}(t) - \hat{\mathcal{G}}_{cs}(t)}{a_S\psi(t\,|\,\hat{\boldsymbol{\Xi}}_{\boldsymbol{Y}}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{Y}})\sqrt{m}} \xrightarrow{\mathcal{D}} N(0,1)\ \left(n\to\infty, \sqrt{n/m} = O(1)\right)$$

上記の式を用いると，

[A9] $$\mathrm{P}\left(\frac{\left|\mathcal{G}_{\mathrm{cs}}(t)-\hat{\mathcal{G}}_{\mathrm{cs}}(t)\right|}{a_S\psi(t\mid\hat{\boldsymbol{\Xi}}_{\boldsymbol{Y}},\hat{\boldsymbol{\Sigma}}_{\boldsymbol{Y}})\sqrt{m}}\leq z_{\alpha/2}\right)\to 1-\alpha\ \left(n\to\infty,\sqrt{n/m}=O(1)\right)$$

となる．[A9] 式を変形することにより [34] 式を得ることができ，定理を証明することができる．