

ノート (Note)

樹種間差および測定誤差を考慮した胸高直径－樹高関係のベイズ推定

伊東 宏樹^{1)*}

キーワード：アロメトリー、階層ベイズモデル、拡張相対成長式、Stan、統計モデル

胸高直径から樹高を推定する関係式はいくつか提案されているが、拡張相対成長式 (Ogawa et al. 1965) もそのためによく使われる。直径 (D) と樹高 (H) についての拡張相対成長式は以下の式で与えられる。

$$\frac{1}{H-H_D} = \frac{1}{aD^h} + \frac{1}{H_{MAX}}$$

ここで、 H_D は樹高測定高、また a および h 、 H_{MAX} はパラメーターであり、とくに H_{MAX} は最大樹高である。

拡張相対成長式のパラメーターは、非線形最小二乗法または最尤法 (島谷 2012) により推定することが可能である。しかし、多樹種のパラメーターを同時推定する場合や、測定誤差を考慮した場合には単純な非線形最小二乗法は適用できず、最尤法でも困難となる。このような場合でも、MCMC 法 (Markov Chain Monte Carlo; マルコフ連鎖モンテカルロ法) を利用したベイズ推定ならばパラメーターの推定ができる可能性がある。今回、新しいベイズ推定ソフトウェアである Stan (Stan Development Team 2014) を使用して、樹種間差および測定誤差を組み込んだ拡張相対成長式のパラメーター推定をおこなうことができたので、その方法を報告する。Stan はハミルトニアンモンテカルロ法を採用することで、従来のベイズ推定ソフトウェアよりも効率的に推定をおこなうとされる。

実際のデータとして、伊東 (2001) にて使用したデータを利用した (表 S1)。このデータは銀閣寺山国有林の 15 樹種の幹から得たものであり、各樹種について 12～17 本の幹、合計 232 本を測定したものである。胸高直径は各幹 1 回の測定だが、樹高は、測定誤差を考慮して 1～4 回の測定をおこなった。

幹 i の樹高の j 回目の測定値を H_{ij} (m)、幹 i の真の樹高から胸高の 1.3 m を引いた値を \bar{H}_i (m) とした。樹高の測定値は、次式のような対数正規分布に従うとした。

ただし、

$$\text{LogNormal}(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma x}} \exp\left(-\frac{1}{2}\left(\frac{\log x - \mu}{\sigma}\right)^2\right) \text{ である。}$$

$$H_{ij} - 1.3 \sim \text{LogNormal}(\log \bar{H}_i, \sigma_H^2)$$

また、幹 i の胸高直径の測定値を D_i (cm)、幹 i の「真の胸高直径」の値を \bar{D}_i (cm) とした。胸高直径の幹ごとの誤差は次式のような正規分布に従うとした。

$$D_i \sim \text{Normal}(\bar{D}_i, \sigma_D^2)$$

今回のデータでは胸高直径は各幹について 1 回しか測定しておらず、測定誤差の推定はできない。ここでいう「真の胸高直径 \bar{D}_i 」とは、推定式へのあてはまりをよくするため、胸高直径を適宜増減させた結果を意味する。このため、幹ごとの誤差 $\bar{D}_i - D_i$ には、胸高直径の測定誤差と幹ごとの推定式からのずれとが含まれる。

拡張相対成長式から以下の関係が導かれる。

$$\log \bar{H}_i = \log a_{s_i} + h_{s_i} \log \bar{D}_i + \log H_{MAX_{s_i}} - \log(a_{s_i} \bar{D}_i^{h_{s_i}} + H_{MAX_{s_i}}) + \varepsilon_i$$

ここで、 s_i は幹 i の樹種であり、 a_{s_i} 、 h_{s_i} 、 $H_{MAX_{s_i}}$ は樹種 s_i のパラメーターである。とくに $H_{MAX_{s_i}}$ は樹種 s_i の最大胸高直径である。また、 ε_i は、拡張相対成長式により推定される樹高からの個体ごとの誤差である。樹種ごとのパラメーター a_{s_i} 、 h_{s_i} 、 $H_{MAX_{s_i}}$ は、それぞれの平均 \bar{a} 、 \bar{h} 、 \bar{H}_{MAX} から、樹種ごとの誤差 $\varepsilon_{a_{s_i}}$ 、 $\varepsilon_{h_{s_i}}$ 、 $\varepsilon_{H_{MAX_{s_i}}}$ を対数スケールでもつとした。個体ごとの誤差 ε_i および樹種ごとの誤差 $\varepsilon_{a_{s_i}}$ 、 $\varepsilon_{h_{s_i}}$ 、 $\varepsilon_{H_{MAX_{s_i}}}$ の事前分布は階層事前分布とし、平均が 0、標準偏差がそれぞれ σ および σ_a 、 σ_h 、 $\sigma_{H_{MAX}}$ の正規分布にしたがうとした。樹種ごとの差を階層化してモデリングすることにより、多樹種についてパラメーターを同時推定することができる。これにより、樹種ごとに個別に推定するよりも安定した結果が得られると考えられる。また、データの少ない部分でも、事前分布 (この場合は樹種間差を含めない平均的な値の分布) に近づくものの、一応の推定が可能となる。

$$\varepsilon_i \sim \text{Normal}(0, \sigma^2)$$

$$\log a_{s_i} = \log \bar{a} + \varepsilon_{a_{s_i}}, \quad \varepsilon_{a_{s_i}} \sim \text{Normal}(0, \sigma_a^2)$$

$$\log h_{s_i} = \log \bar{h} + \varepsilon_{h_{s_i}}, \quad \varepsilon_{h_{s_i}} \sim \text{Normal}(0, \sigma_h^2)$$

$$\log H_{MAX_{s_i}} = \log \bar{H}_{MAX} + \varepsilon_{H_{MAX_{s_i}}}, \quad \varepsilon_{H_{MAX_{s_i}}} \sim \text{Normal}(0, \sigma_{H_{MAX}}^2)$$

Bayesian estimation of tree diameter–height relationships involving species differences and measurement errors
Hiroki ITO^{1)*}

原稿受付：平成 27 年 1 月 30 日 原稿受理：平成 27 年 3 月 11 日

1) 森林総合研究所森林植生研究領域

* 森林総合研究所森林植生研究領域 〒305-8687 茨城県つくば市松の里 1

Department of Forest Vegetation, Forestry and Forest Products Research Institute (FFPRI), 1 Matsunosato, Tsukuba, Ibaraki, 305-8687, Japan;
e-mail: hiroki@ffpri.affrc.go.jp

パラメーター \bar{D} 、 σ_D 、 σ_H 、 σ 、 σ_a 、 σ_h 、 $\sigma_{H_{MAX}}$ の事前分布は $[0, 1000]$ の一様分布とした。また、 $\log \bar{a}$ 、 $\log \bar{h}$ 、 $\log \bar{H}_{MAX}$ 事前分布は平均が 0、標準偏差が 100 の正規分布とした。

Stan のモデリング言語により記述したモデルをリスト S1 に示す。4 本のマルコフ連鎖を用いて、各鎖について 35000 回の繰り返し計算をおこなったが、最初の 5000 回は使用せず、後の 30000 回から 30 回ごとに各パラメーターの事後分布のサンプルを採取した。Gelman-Rubin 統計量と、各マルコフ連鎖の軌跡を目視することにより、収束を判断した。

計算により得られたサンプルから求めた Gelman-Rubin 統計量はすべてのパラメーターについて 1.04 より小さく、また各マルコフ連鎖の軌跡も十分に重なり合っており、計算結果は定常分布に収束していると考えられた。樹種差を含めない平均的な最大樹高の対数 $\log \bar{H}_{MAX}$ の事後平均は 2.83、95% 信用区間は 2.49 ~ 3.23 と推定された (表 S2)。この事後平均の値は樹高 16.9 m 程度に相当する。樹高の測定誤差のパラメーター σ_H の事後分布の平均は 0.03 と推定された。この値は、測定値が真の値の 0.94 ~ 1.06 倍の範囲におさまる確率が 95% であることをしめす。胸高直径の誤差の標準偏差 σ_D の事後分布の平均は 0.50 となった。これは、胸高直径の誤差が ± 0.98 cm の範囲にある確率がおよそ 95% の確率であることに相当する。

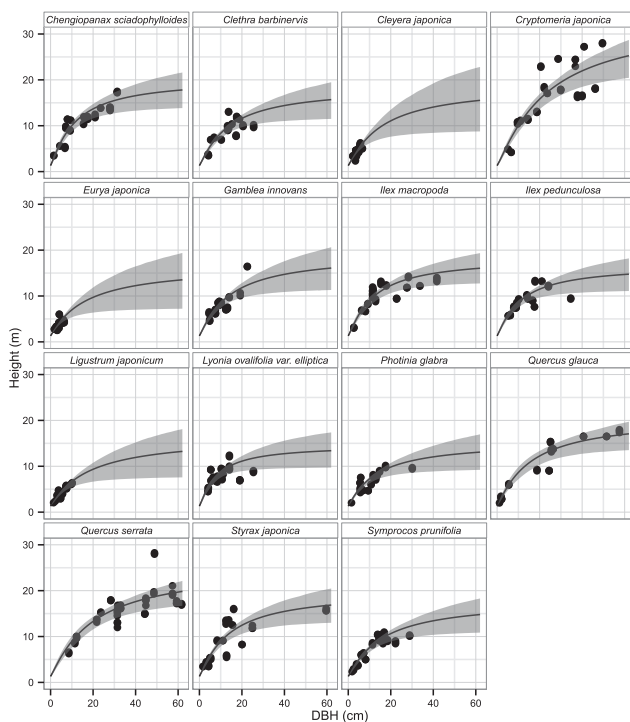


図 1. 各樹種についての胸高直径および樹高の測定値の平均 (点) と、推定された胸高直径–樹高曲線 (実線)、および樹高の 95% 信用区間 (灰色の領域)

樹種ごとのパラメーターの事後分布の平均および 95% 信用区間を表 S3 にしめた。事後平均から予測される各樹種の最大樹高の予測値はスギ (*Cryptomeria japonica*) で最も大きく 31.9 m と推定された。次いで大きかったのはコナラ (*Quercus serrata*) で、23.8 m であった。一方、最も小さかったのはネジギ (*Lyonia ovalifolia* var. *elliptica*) で 13.2 m、次いでカナメモチ (*Photinia glabra*) で 14.0 m であった。各樹種の測定値と、パラメーターの事後平均により予測される胸高直径–樹高曲線と、樹高の 95% 信用区間とを図 1 にしめた。各樹種の最大樹高の事後平均はおおむね妥当な値が得られたと考えられる。ただし、信用区間の幅は広く、かなりの不確実性が認められた。とくに、サカキ (*Cleyera japonica*) やヒサカキ (*Eurya japonica*)、ネズミモチ (*Ligustrum japonicum*) などでは、データが得られていない部分が広く、こうした箇所では事後平均は事前分布に近くなり、信用区間は幅が広がっていると考えられた。

以上の結果をまとめると、各パラメーターの事後平均にはおおむね妥当な値が得られており、拡張相対成長式のモデル化として今回のモデルは妥当であると考えられた。ただし、不確実性を減少させるためには、さらに大きなサンプルサイズや、幅広い樹木サイズからのサンプリングが必要であると考えられた。

引用文献

- 伊東宏樹 (2001) シミュレーションモデルによる広葉樹二次林の遷移の予測—銀閣寺山国有林について—。森林応用研究 10(2), 27–35.
- Ogawa, H., Yoda, K., Ogino, K. and Kira, T. (1965) Comparative ecological studies on three main types of forest vegetation in Thailand II Plant biomass. Nature and life in Southeast Asia 4, 49–80.
- 島谷健一郎 (2012) フィールドデータによる統計モデリングと AIC. 近代科学社, 216pp.
- Stan Development Team (2014) “Stan modeling language user's guide and reference manual, version 2.5.0”, 408pp, <http://mc-stan.org/> (参照 2015-01-29).

補足電子資料

以下はオンライン版のみの掲載となります。

表 S1 測定した幹の樹種および測定値

表 S2 樹種共通パラメーターの事後分布の平均および分位点

表 S3 各樹種のパラメーター a 、 h 、 H_{MAX} の事後分布の平均および分位点

リスト S1 パラメーター推定に使用した Stan モデルファイル

URL : <http://www.ffpri.affrc.go.jp/pubs/bulletin/435/index.html>