

## 国家森林資源調査データに基づく地域レベルの平均蓄積推定 —衛星データを利用した事後層化による推定精度の改善—

西園朋広・田中真哉（森林総研）・北原文章（森林総研四国）・小谷英司（森林総研東北）・  
高橋興明・家原敏郎（森林総研）・齋藤英樹（森林総研九州）

**要旨：**東北地方の一部地域を対象に、衛星データを使って事後層化を行い、国家レベルで設計された森林資源モニタリング調査データから推定した地域レベルの平均蓄積の精度の変化を調べた。ランドサット ETM+データから得た 16 の変数（輝度値、反射率および 4 つの指数）それぞれに基づいて、対象地域を 1～10 の層に分けた。12 の変数で、事後層化後の平均蓄積の精度が、層化前よりも向上した。衛星データを利用した事後層化によって、地域レベルの平均蓄積の推定精度が向上することがわかった。

**キーワード：**事後層化、蓄積推定、ランドサット ETM+データ、国家森林資源調査、森林資源モニタリング調査、

**Abstract:** We investigated changes in precision of estimate for mean stand volume at regional-scale from national forest inventory data by post-stratification using satellite imagery for a part of Tohoku region. Test area was grouped into several (from 1 to 10) strata using each one of 16 variables (digital number and reflectance of bands 1-5 and 7, and 4 spectral indices) obtained from Landsat ETM+ data. For 12 variables, precision of mean stand volume from stratified estimation technique was better than those under the assumption of simple random sampling. We concluded that post-stratification using satellite imagery can increase precision of mean stand volume at regional-scale.

**Key words:** estimate of stand volume, forest resource monitoring survey in Japan, Landsat ETM+ data, national forest inventory, post-stratification

### I はじめに

林野庁の森林資源モニタリング調査（以下、FRMS と記す）は、日本全体で蓄積を推定した場合の誤差率が 3% 以内に入るように設計されている（2）。同調査のように国家レベルで設計された調査のデータから、地域レベルの蓄積を推定すると、小標本のため誤差率が大きくなり、推定精度が低下してしまう。この問題に対し、スウェーデン（5）や米国（3）では衛星データによる事後層化を行い、地域レベルの蓄積推定の精度を高めている。

著者らの知る限り、これまでわが国の国家レベルの森林資源調査において事後層化が適用された事例はなく、わが国の森林について事後層化によって精度向上が見込めるかどうか不明である。本研究では、東北地方の一部地域を対象に、ランドサット ETM+データを使って事後層化を行い、地域レベルの平均蓄積推定の精度の変化を調べ、主として、次の 2 点を明らかにした。①衛星データに基づく事後層化によって推定精度は向上するか？②精度が向上する場合、ランドサット ETM+データから簡便に得られる変数のうち、精度の向上効果に優れた変数は何か？

### II 資料

衛星データと FRMS のデータを資料として用いた。

用いた衛星データは、Landsat/ETM+（パス 107 ロウ 31-34）の 2002 年 6 月に観測された画像である。データの入手先・処理レベルは田中ら（8）と同じである。雲やヘイズの影響で解析に適さない箇所は目視によって除外した。さらに、別途開発中の森林・非森林の分類図（8）を用いて、森林部分のみを抽出した。この抽出した範囲（以



図1 対象地域  
Fig. 1. Test site.

Tomohiro NISHIZONO, Shinya TANAKA (FFPRI, Ibaraki 305-8687), Fumiaki KITAHARA (FFPRI, Shikoku), Eiji KODANI (FFPRI, Tohoku), Tomoaki TAKAHASHI, Toshiro IEHARA (FFPRI), and Hideki SAITO (FFPRI, Kyusyu) Estimation of mean stand volume at regional-scale based on national forest inventory - increase in precision by post-stratification using satellite imagery.

下、範囲Aと記す)を本研究の対象地とした(図1)。すなわち、本研究における母集団は範囲Aの森林である。

用いたFRMSデータは、範囲Aに含まれるプロット(n=1704)の林分材積である。調査の実施年度は1999~2003年である。また、プロットごとに求めた優占構成樹種群も用いた。プロット内の全ての樹種を常緑針葉樹・落葉針葉樹・常緑広葉樹・落葉広葉樹の4樹種群に再分類し、そのなかで断面積割合の最も大きな樹種群を優占構成樹種群とした。

## II 方法

1. 精度の評価 本研究では、平均蓄積の推定精度を評価しようとしている。FRMSのようなサンプリング調査から得られるのは、蓄積の標本平均 $\hat{\mu}$ である。 $\hat{\mu}$ は、標本抽出を繰り返すごとにさまざまな値を取り、ある分布を持つことになる。一般に、サンプリング調査で得られた $\hat{\mu}$ は、母平均 $\mu$ の周りに正規分布することが知られている。この分布の分散の値が小さいほど、抽出標本から得られる平均値のバラツキが小さいことになり、直接はかることのできない母平均を精度よく推定することができる(7)。つまり、精度はサンプリングの手法を繰り返し用いるときに得られる母平均からの偏差の大きさを意味しており(1)、平均推定値の信頼性の指標である。本研究では、以下のような平均値の平均平方誤差(MSE( $\hat{\mu}$ ), 標本平均の分散Var( $\hat{\mu}$ )に等しい)を基本とする指標を用いて、平均蓄積の推定精度を評価した。

$$[1] \quad MSE(\hat{\mu}) = E(\hat{\mu} - \mu)^2 = Var(\hat{\mu})$$

2. 事後層化した場合の統計量の算出 層化抽出法は母集団をいくつかの層に分割し、各層ごとに標本を抽出して、母集団に対しての推定を行う方法である(6)。通常、抽出の前に事前情報に基づいて層を形成することによって、層別あるいは全体の標本数を少なくし、推定の精度を上げることが期待できる。本研究では、抽出前ではなく、抽出後(事後)に補助情報に基づいて層化を行うことによって、標本数は一定のままでも、推定の精度を上げることが期待した。

事後層化後の林分材積の平均値 $\bar{V}_{PS}$ とその分散Var( $\bar{V}_{PS}$ )は、通常の(事前)層化抽出法と同様に次式で求めた(1, 3)。

$$[2] \quad \bar{V}_{PS} = \sum_{h=1}^H W_h \bar{V}_h$$

$$[3] \quad Var(\bar{V}_{PS}) = E(\bar{V}_{PS} - \mu)^2 = \sum_{h=1}^H W_h^2 \hat{\sigma}_h^2 / n_h$$

ただし、 $\bar{V}_h = 1/n_h \sum_{i=1}^{n_h} V_{hi}$ ,  
 $\hat{\sigma}_h^2 = 1/(n_h - 1) \sum_{i=1}^{n_h} (V_{hi} - \bar{V}_h)^2$  である。

ここで、Hは階層の数、 $W_h$ ,  $\bar{V}_h$ ,  $n_h$  および  $V_{hi}$  は階層hにおける重み、平均林分材積、プロット数、i番目のプ

ロットの林分材積であり、 $\mu$ は母集団における平均林分材積である。 $W_h$ は(階層hに含まれるピクセルの数)/(範囲Aの総ピクセル数)として求めた(3)。なお、有限補正項は無視した。

3. 衛星データによる層化 層を形成するためにLANDSAT/ETM+データのバンド1-5および7の輝度値(以下、DN1~DN5, DN7と記す)と反射率(以下、Ref11~Ref15, Ref17と記す)、ならびに反射率から求めた正規化植生指数と反射率をタッセルドキャップ変換して求めたBrightness, Greenness, Wetness(以下、NDVI・TCb・TCg・TCwと記す)の16変数の値を用いた。NDVIとタッセルドキャップ変換値は、植生に関する情報を持つことが期待されるため、蓄積推定に利用される場合がある(3)ことや、教師データを用いずに、反射率のみから簡便に計算できることから検討の対象とした。反射率と4つの指数はErdas Imagine 2011を用いて求めた。大気補正・地形補正は行わなかった。

16変数のそれぞれを補助情報として、範囲Aをいくつかの層に分割した。層の数は1から10までを検討した。ここで、層の数が1の時は、事後層化を行わない事例を意味している。層間の閾値 $y_h$ が、 $y_h = (m_h + m_{h+1})/2$ を満たすように閾値を決定した(5)。ここで、 $m_h$ は補助情報値の小さい順に並べた場合のh番目の階層に属するピクセルの補助情報値の平均値である。閾値は、解析的には得られないので、初期値を与えて、 $y_h$ の残差平方和が最小になるように最適解を求めた。初期値は層ごとのピクセル数が等しくなるように与えた。

FRMSプロットの中心点が含まれる一つのピクセルが所属する層をそのFRMSプロットの所属する層とみなした。なお、Ref11-3が極端に大きかった(例えば、Ref11 > 0.145)3つのプロットを外れ値と見なして除外し、最終的に1701プロットを解析の対象とした。

4. 精度に対して事後層化が与える影響の評価 平均蓄積の分散((1), (3)式)から、標準誤差SE, 誤差率E, 相対効率REの3つの指標(1)を算出して、事後層化が精度へ与える影響を評価した。それぞれ次式で求めた。

$$[4] \quad SE = \sqrt{Var(\bar{V}_{PS})}$$

$$[5] \quad E = t_{0.05} \cdot SE / \bar{V}_{PS}$$

$$[6] \quad RE = Var(\bar{V}_{SYS}) / Var(\bar{V}_{PS})$$

ここで、 $t_{0.05}$ は危険率5%におけるt値であり、 $Var(\bar{V}_{SYS})$ は、層化しない場合(単純系統抽出法)の平均の分散である。後者は、(3)式においてH=1とした場合と等しい。SEとEが小さいほど、REが1より大きいほど、事後層化による精度の改善効果が大きいといえる。階層数1~10について、3つの指標を算出した。

5. 衛星データと蓄積との相関解析 McRoberts et al. (3)によると、層化に利用する変数と蓄積との関連が深いほど、推定精度の改善が期待できる。16変数と蓄積との間の相関係数を計算し、事後層化による改善効果との関連を検討した。さらに、事後層化による精度改善の仕組みを把握する参考として、優占樹種群別に16変数と蓄積との間の相関係数を計算した。また、16変数と蓄積について、優占樹種群間で平均値に差があるかどうかをスチール・ドゥアス検定で調べた。ただし、範囲Aにおいて常緑広葉樹が優占したのは1プロットのみだったので、解析から除外し、最終的に3樹種群間で比較した。

#### IV 結果と考察

1. 層化による推定精度の改善効果 12変数で層化による改善効果が認められた(表1, 図2)。しかし、DN1・DN3・Ref11・Ref13については、層化による改善効果は認められず、事後層化によってSEとEは増加し、REは1以下であった(図2: SEは省略)。改善効果が認められた12変数のうちNDVIを除いた11変数においては、階層数が増加するほどSEとEが減少し、REが大きくなる傾向があった。しかし、その変化率は漸減傾向にあって、頭打ちがみられた(図2)。REが最大値に達したときの階層数は変数ごとに異なっていた(表1)。NDVIは他の変数と比べて不規則に変化した。

2. ランドサットデータの種類と改善効果の関係 改善効果は、DN5, Ref15, およびTCwで大きく、REの最大値(以下、maxREと記す)はそれぞれ、1.128, 1.136, 1.131であった(表1)。この3変数では、全プロットにおける変数と林分材積との間の相関係数の絶対値(以下、 $|r_{vol, All}|$ と記す)が他の変数と比べて大きかった(それぞれ、0.408, 0.413および0.410)。

16変数におけるmaxREと $|r_{vol, All}|$ の間には、強い正の相関が認められた( $r=0.938$ ,  $p<0.001$ )。この結果は、層化に利用する変数と蓄積との関連が深いほど、推定精度の改善が期待できるという知見(3)に合致している。

$|r_{vol, All}|$ は、全ての変数において、常緑針葉樹林と落葉広葉樹林における相関係数の絶対値(以下、 $|r_{vol, EC}|$ および $|r_{vol, DB}|$ と記す)よりも大きく、大多数の変数において落葉針葉樹林における相関係数の絶対値(以下、 $|r_{vol, DC}|$ と記す)よりも小さかった(表1)。他の樹種群では、複数の優占樹種が存在したのに対して、落葉針葉樹林の優占樹種はカラマツのみであったことから、光の反射に影響する樹冠の特性が樹種群内で均一であったことが後者の原因かもしれない。 $|r_{vol, All}|$ が大きいほど、 $|r_{vol, EC}|$ と $|r_{vol, DC}|$ は大きかったが( $r=0.975$ ,  $0.945$ , 共

に $p<0.01$ )、 $|r_{vol, All}|$ と $|r_{vol, DB}|$ の間には関連はなかった( $r=-0.336$ ,  $p=0.204$ )。よって、常緑・落葉針葉樹林における相関の強弱は全プロットの相関の強弱に影響していたが、落葉広葉樹林における相関の強弱は影響していないといえる。

$|r_{vol, All}|$ が大きい( $>0.27$ )変数で、3つの優占樹種群間全てにおいて平均値に差が認められる場合が多かった(表2)。このときの各変数の平均値の絶対値は、常緑針葉樹<落葉針葉樹<落葉広葉樹の順に大きかった。林分材積の平均値は、常緑針葉樹>落葉針葉樹>落葉広葉樹の順に小さかった。よって、数直線上に常緑針葉樹・落葉針葉樹・落葉広葉樹の順(あるいは逆順)で樹種群が並ぶような変数において、樹種群間での林分材積のサイズ順位を反映して、全プロットの相関が大きくなったと考える。

以上から、本研究で対象とした変数の中で改善効果の高かった変数の特徴は次の2点だと考える。①針葉樹林において蓄積の樹種群内差異を反映した情報を持っていた。②落葉広葉樹と針葉樹との間の樹種群間差異を反映した情報を持っていた。①と②の効果が合わさって、全プロットを対象にしたときに蓄積との間に比較的強い相関を持ち、改善効果が大きかったといえる。

#### V 今後の課題

本研究では、わが国の国家森林資源調査において、衛星データを事後層化のための情報源とすることが有用であることがわかった。米国やスウェーデンでは、中分解能衛星データから作成した蓄積マップを事後層化に利用して、蓄積の推定精度を向上させている。また、森林タイプ分類図の利用もみられる。これらの事例のRE(1.25~3.72)は本研究のそれよりも大きく、推定精度向上の効果が大きい(4)。地域レベルの資源統計量の精度向上を指向する観点から、より効果の高い補助情報源が必要である。そのために、衛星データを利用した蓄積推定図・森林タイプ分類図の作成に関する研究の進展が必要である。また、本研究と植生条件の異なる地域において同様の解析を行い、地域特性の有無を検討することが、わが国全体で統計情報を整備する上で重要だと考える。

本研究は農林水産省プロジェクト研究「農林水産分野における地球温暖化対策のための緩和及び適応技術の開発」の小課題「リモートセンシングと地上観測による林分構造の広域評価手法の開発」の一環として実施した。

#### 引用文献

- (1) COCHRAN, W.G.(1977) Sampling techniques. 3rd ed.

Wiley, New York.

- (2) 家原敏郎 (1999) 山林 1384 : 54-61.
- (3) McROBERTS et al. (2006) Can. J. For. Res. 36: 2968-2980.
- (4) McROBERTS & TOMPP0 (2007) Remote. Sense. Environ. 110: 412-419.
- (5) NILSSON et al. (2005) Proc. of ForestSat 2005, Borås,

Sweden, 31 May-3 June, Report 8a (pp. 22-26)

- (6) 西沢正久 (1965) 改訂森林測定法. 302pp., 地球出版, 東京.
- (7) 林野庁計画課 (2009) 森林資源モニタリング調査実施マニュアル. 97pp., 林野庁, 東京.
- (8) 田中真哉ら (2011) 関東森林研究 62 : 79-82.

表1. 事後層化によって得られた精度指標の最適値・最適階層数・最適条件での平均蓄積、およびFRMSプロットにおけるランドサットデータと林分材積との相関係数

Table 1. Optimum precision index, optimum number of stratification, estimate of mean stand volume under the optimum condition, and correlation coefficient between Landsat data used for stratification and stand volume at inventory plot.

Landsat data used for stratification	Error ratio (E)			Standard error of mean (SE) and relative efficiency (RE)				Correlation coefficient between Landsat data and stand volume*2							
	Minimum E (%)	Number of strata	Mean stand volume (m <sup>3</sup> /ha)	Minimum SE (m <sup>3</sup> /ha)	Maximum RE	Number of strata	Mean stand volume (m <sup>3</sup> /ha)	All plots		Evergreen coniferous		Deciduous coniferous		Deciduous broadleaved	
								r	p value	r	p value	r	p value	r	p value
DN1	3.233	1	217.98	3.593	1.000	1	217.98	0.054	0.03	0.046	0.23	0.098	0.44	-0.122	<0.01
DN2	3.117	4	213.07	3.374	1.065	7	212.05	-0.326	<0.01	-0.242	<0.01	-0.308	0.01	-0.257	<0.01
DN3	3.233	1	217.98	3.593	1.000	1	217.98	-0.149	<0.01	-0.138	<0.01	-0.191	0.13	-0.239	<0.01
DN4	2.937	9	216.37	3.240	1.109	9	216.37	-0.368	<0.01	-0.233	<0.01	-0.382	<0.01	0.011	0.73
DN5	2.917	6	214.11	3.185	1.128	6	214.11	-0.408	<0.01	-0.267	<0.01	-0.479	<0.01	-0.084	<0.01
DN7	2.961	9	213.32	3.218	1.116	6	212.92	-0.388	<0.01	-0.252	<0.01	-0.449	<0.01	-0.164	<0.01
Ref11	3.233	1	217.98	3.593	1.000	1	217.98	0.050	0.04	0.048	0.20	0.111	0.38	-0.133	<0.01
Ref12	3.104	5	211.62	3.349	1.073	5	211.62	-0.337	<0.01	-0.247	<0.01	-0.309	0.01	-0.277	<0.01
Ref13	3.233	1	217.98	3.583	1.003	4	215.30	-0.154	<0.01	-0.134	<0.01	-0.214	0.09	-0.253	<0.01
Ref14	2.934	6	216.73	3.242	1.108	6	216.73	-0.371	<0.01	-0.236	<0.01	-0.388	<0.01	0.004	0.90
Ref15	2.903	9	213.64	3.162	1.136	9	213.64	-0.413	<0.01	-0.268	<0.01	-0.486	<0.01	-0.100	<0.01
Ref17	2.961	10	212.47	3.207	1.120	10	212.47	-0.390	<0.01	-0.247	<0.01	-0.448	<0.01	-0.179	<0.01
NDVI	3.122	4	220.24	3.506	1.025	4	220.24	-0.271	<0.01	-0.147	<0.01	-0.195	0.12	0.162	<0.01
TCb	2.922	8	214.34	3.193	1.125	8	214.34	-0.393	<0.01	-0.254	<0.01	-0.420	<0.01	-0.047	0.15
TCg	3.009	8	219.62	3.364	1.068	7	219.08	-0.339	<0.01	-0.206	<0.01	-0.340	<0.01	0.059	0.07
TCw	2.903	10	214.58	3.176	1.131	10	214.58	0.410	<0.01	0.265	<0.01	0.500	<0.01	0.114	<0.01

\*1 DN1-5, 7 and Ref11-5, 7 are digital number and reflectance of bands 1-5, 7. NDVI, TCb, TCg, and TCw are normalized difference vegetation index and tasseled cap brightness, greenness, and wetness.

\*2 Sample sizes for all plots, evergreen coniferous plots, deciduous coniferous plots, and deciduous broadleaved plots were 1701, 693, 64, and 943.

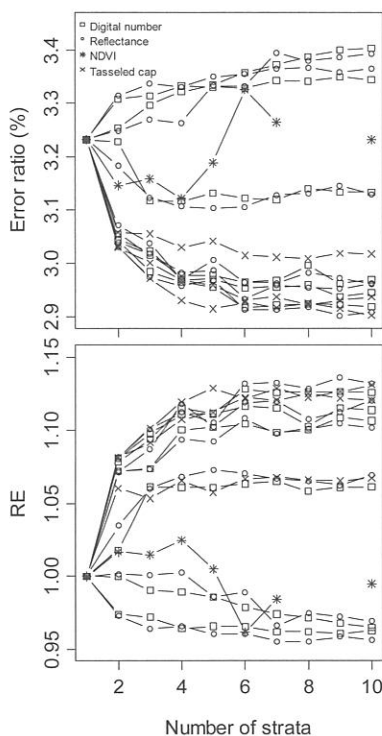


図2. 階層数を変化させた場合の誤差率 (上) と相対効率 (下) の推移

Fig. 2. Changes in error ratio and relative efficiency with increasing number of strata.

表2. FRMSプロットにおけるランドサットデータ・林分材積の平均値の優占樹種群間での比較

Table 2. Multiple comparison between dominant species groups of plot for Landsat data and stand volume at inventory plots.

Variable	Statistics for all plots*1		Mean for different dominant species groups*2			Number of significantly different groups
	Mean	Coefficient of variance (%)	Evergreen coniferous	Deciduous coniferous	Deciduous broadleaved	
DN1	72.614	6.056	73.530 <sup>a</sup>	70.797 <sup>b</sup>	72.065 <sup>b</sup>	2
DN2	59.127	7.795	57.703 <sup>a</sup>	57.547 <sup>a</sup>	60.281 <sup>b</sup>	2
DN3	41.305	12.505	41.433	40.594	41.260	1
DN4	112.581	23.808	92.333 <sup>a</sup>	108.422 <sup>b</sup>	127.743 <sup>c</sup>	3
DN5	84.391	24.629	69.384 <sup>a</sup>	76.766 <sup>b</sup>	95.937 <sup>c</sup>	3
DN7	39.524	22.715	34.326 <sup>a</sup>	36.422 <sup>b</sup>	43.555 <sup>c</sup>	3
Ref11	0.091	6.640	0.092 <sup>a</sup>	0.089 <sup>b</sup>	0.090 <sup>b</sup>	2
Ref12	0.079	9.007	0.077 <sup>a</sup>	0.077 <sup>a</sup>	0.081 <sup>b</sup>	2
Ref13	0.048	15.497	0.048	0.047	0.047	1
Ref14	0.355	25.060	0.288 <sup>a</sup>	0.342 <sup>b</sup>	0.405 <sup>c</sup>	3
Ref15	0.152	27.309	0.122 <sup>a</sup>	0.138 <sup>b</sup>	0.175 <sup>c</sup>	3
Ref17	0.060	28.663	0.050 <sup>a</sup>	0.054 <sup>b</sup>	0.068 <sup>c</sup>	3
NDVI	0.752	9.003	0.706 <sup>a</sup>	0.757 <sup>b</sup>	0.785 <sup>c</sup>	3
TCb	0.374	19.920	0.318 <sup>a</sup>	0.359 <sup>b</sup>	0.416 <sup>c</sup>	3
TCg	0.148	39.098	0.104 <sup>a</sup>	0.143 <sup>b</sup>	0.180 <sup>c</sup>	3
TCw	-0.080	-42.658	-0.056 <sup>a</sup>	-0.068 <sup>b</sup>	-0.098 <sup>c</sup>	3
Stand Volume	217.879	67.980	294.734 <sup>a</sup>	211.550 <sup>b</sup>	161.828 <sup>c</sup>	3

\*1 Sample size was 1700, because one evergreen broadleaved plot was excluded from analysis.

\*2 Different letters indicate significant difference in mean values between dominant species groups of plots ( $p < 0.05$ , Steel-Dwass test).