

# Kapos分類ごとのMGCIの 区間推定値

横浜市立大学大学院データサイエンス研究科教授

佐藤彰洋

2021年2月13日

# FAO試算値 おさらい

日本におけるKapos山地分類#2(3500m-4500m)は富士山山頂付近のみ。JAXA土地利用土地被覆分類によると45ピクセル(約1.96km<sup>2</sup>)が該当し、そのいずれもが裸地(JAXA土地分類#10)である。そのため、**森林、農地は0%**である。Kapos山地分類#2のMGCIは**数%程度**とするべき。

Year 2015

Kapos mountain class

IPCC class	1	2	3	4	5	6
Forest	0.00	20.00	58.89	89.37	94.16	91.10
Cropland	0.00	66.67	33.57	6.34	3.01	4.94
Grassland	0.00	0.00	6.95	3.57	2.03	3.01
Wetland	0.00	0.00	0.35	0.40	0.48	0.08
Settlement	0.00	0.00	0.00	0.00	0.02	0.51
Other land	0.00	13.33	0.24	0.32	0.28	0.36
MGCI	0.00	86.67	99.76	99.68	99.69	99.13
MGCI all			99.23			

今回のターゲット

日本におけるKapos山地分類#3(2500m-3500m)は高地山岳地帯であり、裸地の割合が極めて高い。そのため、山地植生被覆指数は**41.34%**と低い。Kapos山地分類#3における**水田や畑地**は極めて少ないため、現状のFAO試算より**下方修正**が必要である。

# 誤り率を考慮に入れた 95%信頼区間推定値

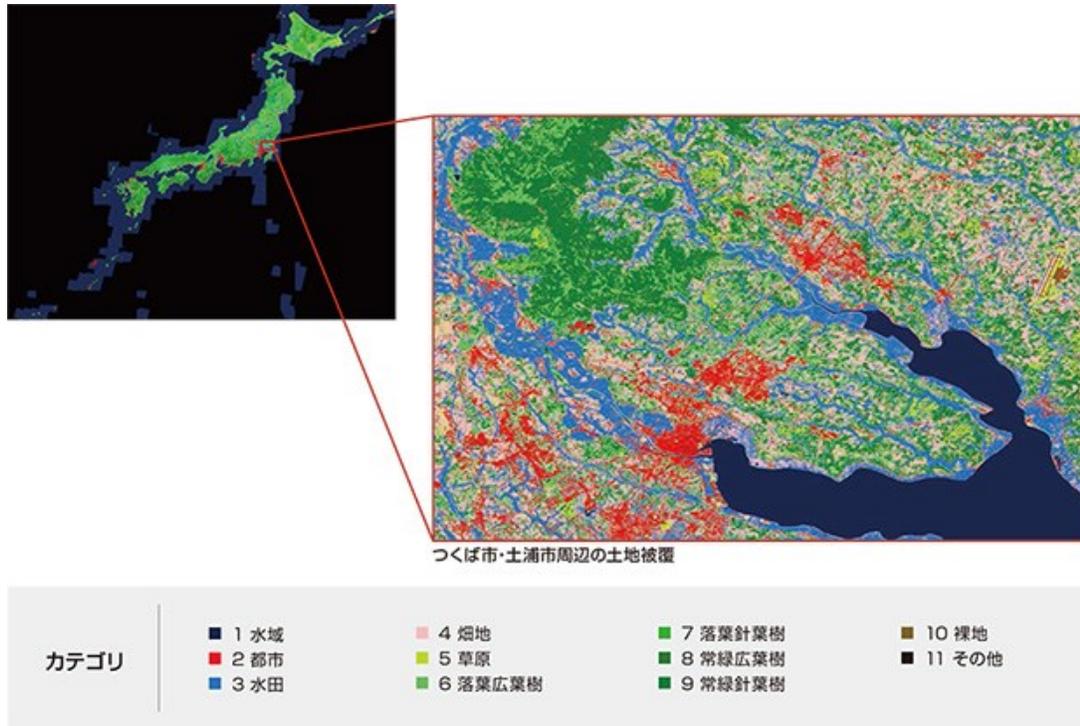
2006-2011 MGCI(全体) 0.996223 (0.885870以上)  
2014-2016 MGCI(全体) 0.988062 (0.867229以上)

年	解像度	指数推計値	グリッド数	指数推計値 95%信頼区間 (下限)	指数推計値 95%信頼区間 (上限)	R(1 1)	R(1 1)95% 信頼区間 (下限)	R(1 1)95% 信頼区間 (上限)	R(1 0)	R(1 0)95% 信頼区間 (下限)	R(1 0)95% 信頼区間 (上 限)	補正值	補正值95% 信頼区間 (下限)	補正值95% 信頼区間 (上限)
2015	300m	0.9907	1921741	0.990564287	0.990835713									
2014- 2016	250m	0.9865	2767043	0.986364024	0.986635976	0.99427 7539	0.9886855 97	0.9998694 82	0.27272727 3	0.18499661	0.360457935	0.989221	0.739504	1.238938
同上	100m	0.9867	17294515	0.986646009	0.986753991	0.9957	0.9909	1.0006	0.12121212 1	0.05692062 6	0.185503617	0.989699	0.837873	1.141525
同上	30m	0.9862	281131587	0.986186363	0.986213637	0.99713 877	0.9931789 84	1.0010985 55	0.08081	0.02712	0.1345	0.988062	0.867299	1.108826
2010	300m	0.9925	1921739	0.992378015	0.992621985									
2006- 2011	250m	0.9857	2767043	0.985560109	0.985839891	0.97641 5094	0.9559872 54	0.9968429 35	0.26436781 6	0.17169953 7	0.357036095	1.013040	0.721796	1.304283
同上	100m	0.9837	17294515	0.98364032	0.98375968	0.98584 9057	0.9699494 62	1.0017486 51	0.05747126 4	0.00856447 9	0.10637805	0.997685	0.875297	1.120074
同上	50m	0.9823	101205994	0.98227431	0.98232569	0.98585	0.9699494 62	1.0017486 51	0.04598	0.00197	0.0899	0.996223	0.885870	1.106576

# 方針

- Kapos#2,#3は対象エリアが狭く中間発表時点で、該当箇所を全て調べることでFAO試算値が過大評価であることが判明した
- Kapos#4,#5,#6は対象エリアが広いため、標本数は十分あるが、土地被覆分類図作成時の分類誤りの影響を大きく受けている
- 方針① JAXA2014年-2016年土地利用土地被覆7.5秒度(250m)グリッドデータKapos#4,#5,#6に対するMGCI推計値を算出するとともに、分類誤りを考慮した補正值の95%信頼区間を算出してみる
- 方針② JAXA2014年-2016年土地利用土地被覆7.5秒度(250m)グリッドデータの植生被覆分類（農地、森林、草地）グリッドを中心に分類誤りを地上観測データとの突合せにより測定することで、95%信頼区間幅を狭くすることを目指す

# JAXA高解像度土地利用土地被覆図



- 最小データは10mメッシュ(0.3秒度×0.3秒度角グリッド)
- 250mグリッドデータ(7.5秒度×7.5秒度角グリッド)
- 水域、都市、水田、畑地、草原、落葉広葉樹、落葉針葉樹、常緑広葉樹、常緑針葉樹、裸地、その他の分類で土地利用・土地被覆分類ラベルがデータ化されている

URL [https://www.eorc.jaxa.jp/ALOS/lulc/lulc\\_jindex.htm](https://www.eorc.jaxa.jp/ALOS/lulc/lulc_jindex.htm)

# 分類ラベル

- #0: 未分類 (Unclassified)
- #1: 水域 (Water)
- #2: 都市 (Urban and built-up)
- #3: 水田 (Rice paddy)
- #4: 畑地 (Crops)
- #5: 草地 (Grassland)
- #6: 落葉広葉樹 (DBF)
- #7: 落葉針葉樹 (DNF)
- #8: 常緑広葉樹 (EBF)
- #9: 常緑針葉樹 (ENF)
- #10: 裸地 (Bare land)
- #11: 雪氷 (Snow and ice)
- #253: その他 (Other)
- #255: データなし (No data)

**評価用データ取得FTPサイト：**

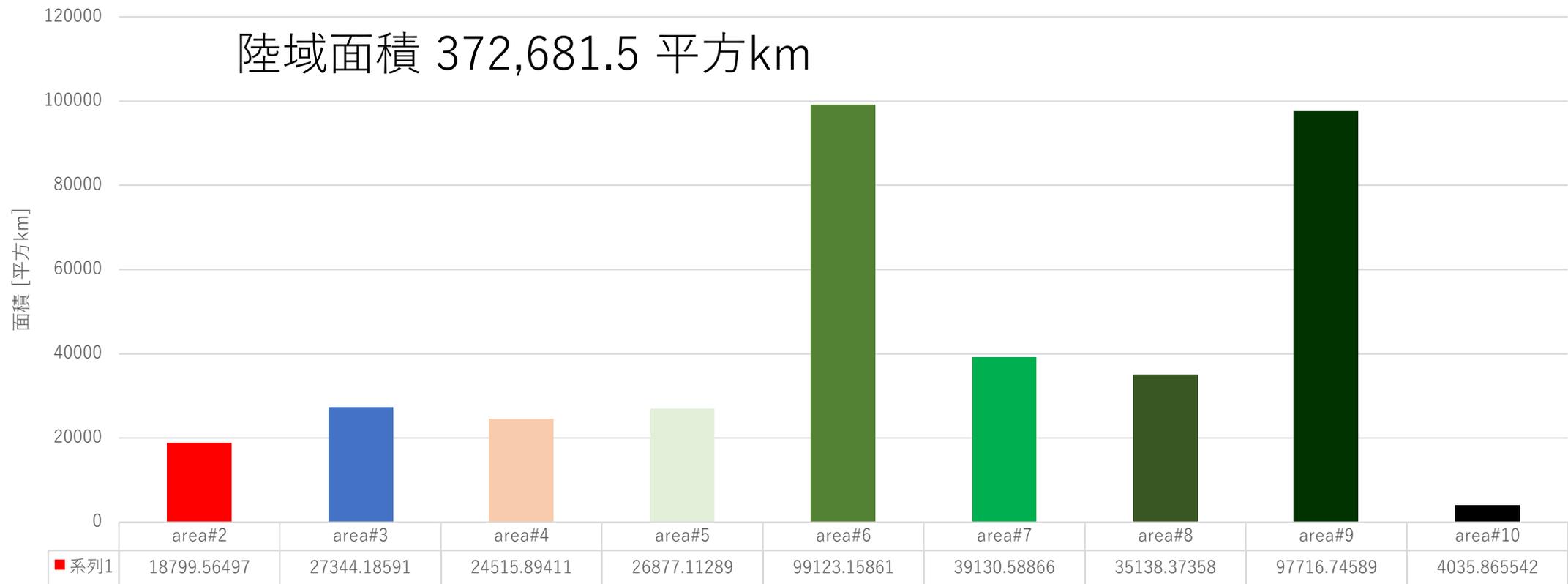
<ftp://ftp.eorc.jaxa.jp/pub/EACH/ps1/common/sdgs>

# 日本全体

- #2: 都市 (Urban and built-up)
- #3: 水田 (Rice paddy)
- #4: 畑地 (Crops)
- #5: 草地 (Grassland)
- #6: 落葉広葉樹 (DBF)

- #7: 落葉針葉樹 (DNF)
- #8: 常緑広葉樹 (EBF)
- #9: 常緑針葉樹 (ENF)
- #10: 裸地 (Bare land)

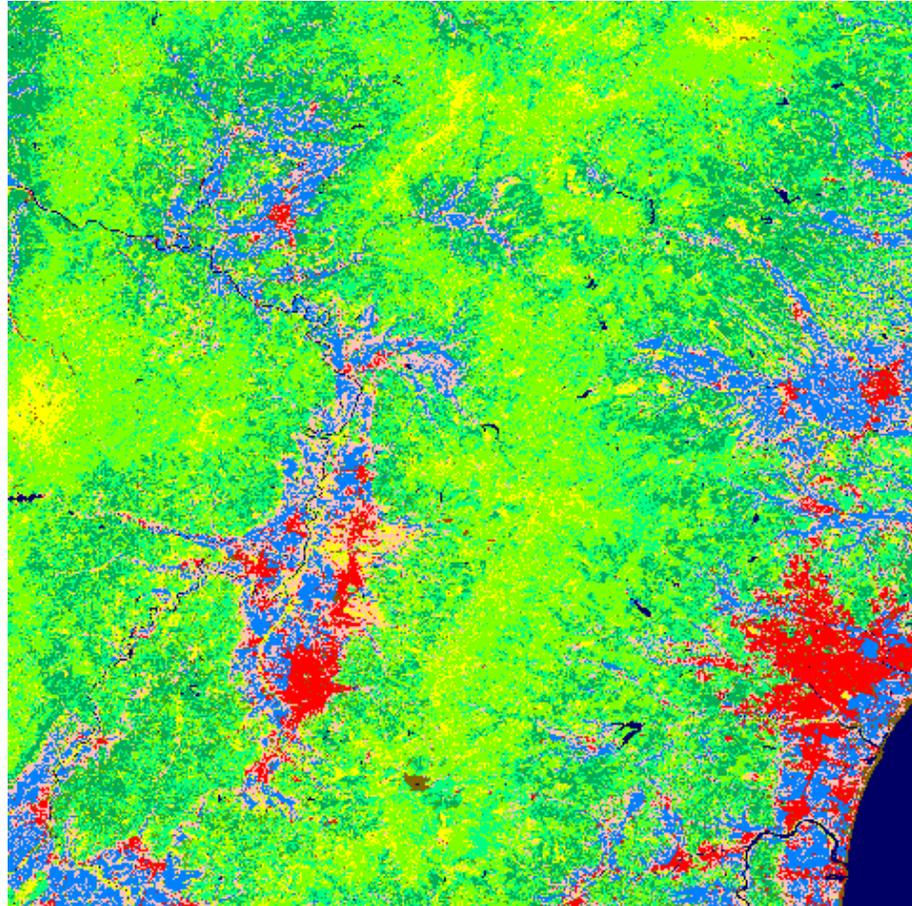
土地利用土地分類面積



<https://www.gsi.go.jp/KOKUJYOHU/MENCHO-title.htm>

令和2年7月1日時点の国土の面積(面積調)は、377,976.94平方km→ 7,000平方km程度が陸水域とみられる<sup>6</sup>

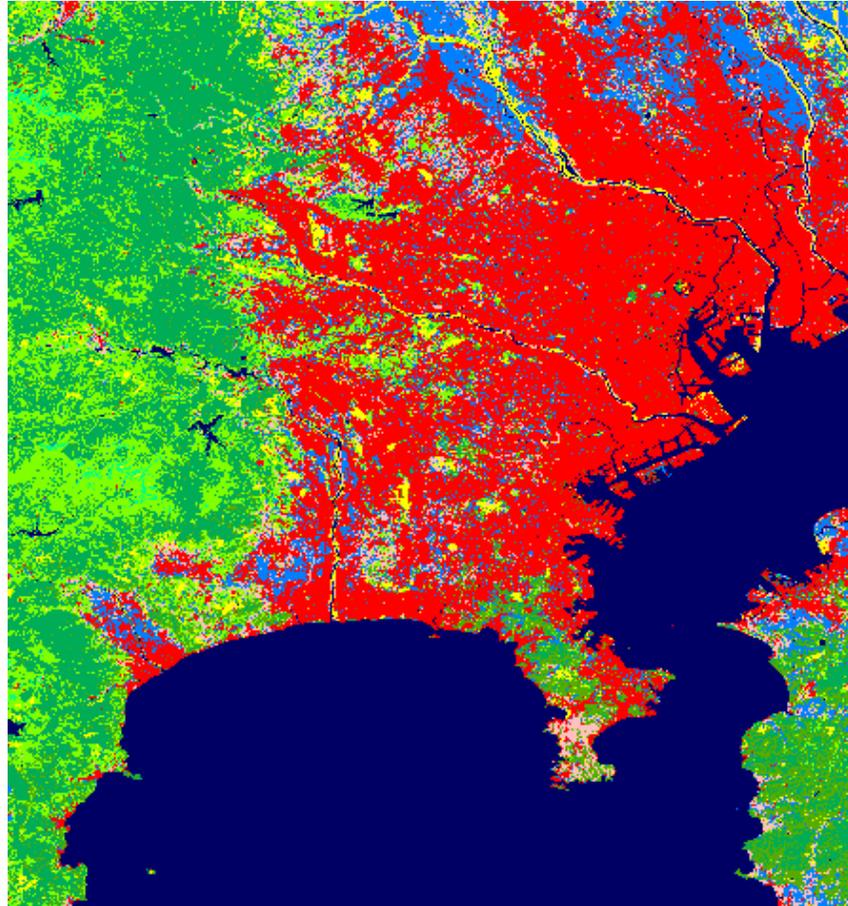
# 宮城県周辺



- 緯度経度 (N38 E140)
- LC\_N38E140\_250m.tif

分類	グリッド数	面積 (km <sup>2</sup> )
#1	4451	187.837
#2	9175	386.7153
#3	20627	867.58
#4	21792	916.9372
#5	18587	781.1932
#6	71668	3010.004
#7	44963	1889.013
#8	6337	266.5762
#9	31026	1302.644
#10	1774	74.73248
合計	230400	9683.232

# 神奈川県周辺



- 緯度経度 (N35 E139)
- LC\_N35E139\_250m.tif

分類	グリッド数	面積 (km <sup>2</sup> )
#1	67492	2957.709
#2	61886	2698.811
#3	16986	740.4813
#4	8923	389.2872
#5	5563	242.6237
#6	21548	940.6618
#7	1183	51.56965
#8	9969	436.6348
#9	35146	1534.211
#10	1704	74.28568
合計	230400	10066.28

$$\sigma_i^2 = \hat{g}_i(1 - \hat{g}_i)/N_i$$

# データからの確認

95%信頼区間： $\hat{g}_i \pm 1.96\sigma_i$

99%信頼区間： $\hat{g}_i \pm 2.58\sigma_i$

- ピクセル数を基にMGCIと95%信頼区間を算出

使用データ

- JAXA 2014年-2016年土地利用土地被覆7.5秒度×7.5秒度(250m)グリッドデータ
- USGS GlobalMountainsK1Classes (k1classes.tif) 7.5秒度×7.5秒度(250m)グリッドデータ

		Kapos1	Kapos2	Kapos3	Kapos4	Kapos5	Kapos6	Kapos0	total
水域	#1	0	0	2	129	1098	15613	1304	18146
都市	#2	0	0	3	116	319	15030	430	15898
水田	#3	0	0	19	354	1529	55270	2391	59563
畑地	#4	0	0	82	698	2675	51677	1198	56330
草地	#5	0	0	886	15567	25782	161184	1584	205003
落葉広葉樹	#6	0	0	264	44217	220335	1181672	10334	1456822
落葉針葉樹	#7	0	0	1109	56596	142864	434766	3826	639161
常緑広葉樹	#8	0	0	424	10400	13561	264794	2669	291848
常緑針葉樹	#9	0	0	185	38794	100391	1248124	10341	1397835
裸地	#10	0	45	4189	5855	3403	10675	189	24356
	Total	0	45	7163	172726	511957	3438805	34266	4164962
	MCGI(%)		0	41.4491135	96.4684	99.05851	98.79848	94.38802	98.59783
	分散		0	9.88296E-05	0.000158	0.000206	2.39E-05	0.000491	1.69E-05
	標準偏差		0	0.009941305	0.012576	0.014336	0.00489	0.022155	0.004109
	95%信頼区間下限		0	41.42962854	96.44375	99.03042	98.78889	94.3446	98.58977
	95%信頼区間上限		0	41.46859846	96.49304	99.08661	98.80806	94.43145	98.60588

# 山地植生被覆指数の推定方法

山地分類*i*領域内の観測点*t*で植生グリッドなら $X_i(t)=1$ 非植生グリッドなら $X_i(t)=0$ とする。山地分類*i*に分類されるグリッド総数 $N_i$ とする。

$$\hat{g}_i = \frac{\sum_{t=1}^{N_i} X_i(t)}{N_i}$$

は推定すべき、真の山地植生被覆指数(MGCI)の値である母比率 $p_i$ の標本比率である。 **$R(a|b)$ を実際(validated)は $b=\{\text{植生, 非植生}\}$ であるが、データ上での分類ラベル(classified)は $a=\{\text{植生, 非植生}\}$ である条件付き確率として定義する。**

いま、 $X_i(t)$ を確率変数とすると、

$$X_i = \begin{cases} 1 & \text{w.p. } R(1|1)p_i + R(1|0)(1 - p_i) \\ 0 & \text{w.p. } R(0|0)(1 - p_i) + R(0|1)p_i \end{cases}$$

となる。

$R(a|b)$ のコンフュージョン  
マトリックスからの推計  
は標本比率なので95%信  
頼区間がある

# $R(a|b)$ の定義

$$R(0|0) = \frac{TN}{FP + TN}, R(1|0) = \frac{FP}{FP + TN}$$

$$R(0|1) = \frac{FN}{TP + FN}, R(1|1) = \frac{TP}{TP + FN}$$

- 真のMGCI母比率を $p_i$ とし、 $R(a|b)$ を実際(validated)は $b=\{\text{植生}, \text{非植生}\}$ であるが、データ分類(classified)上は $a=\{\text{植生}, \text{非植生}\}$ である条件付き確率として定義する。

		validated $b$	
		植生 (X=1)	非植生 (X=0)
classified $a$	植生 (X=1)	$TP$	$FP$
	非植生 (X=0)	$FN$	$TN$

条件付き確率：
$$R(a|b) = \frac{p_{AB}(a, b)}{p_B(b)}$$

$$p_{AB}(a, b) = R(a|b)p_B(b)$$

$$p_A(a) = \sum_b p_{AB}(a, b) = \sum_b R(a|b)p_B(b)$$

$$p_A(1) = \sum_{b=0,1} R(1|b)p_B(b) = R(1|1)p_B(1) + R(1|0)p_B(0)$$

$$p_A(0) = \sum_{b=0,1} R(0|b)p_B(b) = R(0|0)p_B(0) + R(0|1)p_B(1)$$

ここで、 $p_B(1) = p_i, p_B(0) = 1 - p_i$ である。

# 山地植生被覆指数推定値の誤差評価

$$Z_i = \sum_{t=1}^{N_i} X_i(t)$$

とすると、 $Z_i$ は二項分布 $B(N_i, R(1|1)p_i + R(1|0)(1 - p_i))$ に従う。

更に、 $N_i$ が十分大きいと、 $Z_i$ は

平均: $N_i(R(1|1)p_i + R(1|0)(1 - p_i))$ 、

分散: $N_i(R(1|1)p_i + R(1|0)(1 - p_i))(R(0|0)(1 - p_i) + R(0|1)p_i)$

の正規分布により近似できる。よって、山地植生被覆指数推定値

$$\hat{g}_i = \frac{\sum_{t=1}^{N_i} X_i(t)}{N_i}$$

とすると、 $\hat{g}_i$ は

平均: $R(1|1)p_i + R(1|0)(1 - p_i)$ 、

分散: $(R(1|1)p_i + R(1|0)(1 - p_i))(R(0|0)(1 - p_i) + R(0|1)p_i)/N_i$

の正規分布で近似できる。

# 分類誤りが指数推定値に及ぼす影響

- 指数推定値の平均値

$$E[\hat{g}_i] = R(1|1)p_i + R(1|0)(1 - p_i)$$

- 植生/非植生に対する判定誤りがあると指数推定値は真の母指数 $p_i$ よりも減少または増加する
- メッシュにおける植生/非植生の分類判定率が判明している場合は、推定される指数推定値を次の式で補正するのが良いかもしれない

$$p_i = \frac{E[\hat{g}_i] - R(1|0)}{R(1|1) - R(1|0)}$$

# 四則演算（＋－×÷）に対する 誤差伝播公式

- 加算誤差の伝播

$$(a \pm \delta a) + (b \pm \delta b) = (a + b) \pm (\delta a + \delta b)$$

- 減算誤差の伝播

$$(a \pm \delta a) - (b \pm \delta b) = (a - b) \pm (\delta a + \delta b)$$

- 乗算誤差の伝播

$$(a \pm \delta a)(b \pm \delta b) = ab \left( 1 \pm \left( \frac{\delta a}{a} + \frac{\delta b}{b} \right) \right)$$

- 除算誤差の伝播

$$\frac{a \pm \delta a}{b \pm \delta b} = \frac{a}{b} \left( 1 \pm \left( \frac{\delta a}{a} + \frac{\delta b}{b} \right) \right)$$

# 母比率の補正の誤差伝播

$$p_i = \frac{E[\hat{g}_i] - R(1|0)}{R(1|1) - R(1|0)}$$

の誤差伝播について考える。

四則演算に対する誤差の伝播公式より

$$a = E[\hat{g}_i], b = R(1|0), c = R(1|1)$$

$\delta a, \delta b, \delta c$ をそれぞれ $E[\hat{g}_i], R(1|0), R(1|1)$ の95%信頼区間幅の半分とすると、母比率 $p_i$ の信頼区間は以下で与えられる。

$$\frac{a-b}{c-b} \left( 1 - \left( \frac{\delta a + \delta b}{a-b} + \frac{\delta c + \delta b}{c-b} \right) \right) \leq p_i \leq \frac{a-b}{c-b} \left( 1 + \left( \frac{\delta a + \delta b}{a-b} + \frac{\delta c + \delta b}{c-b} \right) \right)$$

# 今後の課題

## →分類誤りの精度検証

- $R(a|b)$ を実際(validated)は $b=\{\text{植生}, \text{非植生}\}$ であるが、データ分類(classified)上は $a=\{\text{植生}, \text{非植生}\}$ である条件付き確率として定義する。

		validated $b$	
		植生 (X=1)	非植生 (X=0)
classified $a$	植生 (X=1)	$TP$	$FP$
	非植生 (X=0)	$FN$	$TN$

- $R(a|b)$ のコンフュージョンマトリックスからの推計は標本比率なので95%信頼区間がある

$$R(0|0) = \frac{TN}{FP + TN}, R(1|0) = \frac{FP}{FP + TN}$$
$$R(0|1) = \frac{FN}{TP + FN}, R(1|1) = \frac{TP}{TP + FN}$$

$$\sigma^2 = \hat{q}(1 - \hat{q})/N$$

$$95\% \text{信頼区間} : \hat{q} \pm 1.96\sigma$$

$$99\% \text{信頼区間} : \hat{q} \pm 2.58\sigma$$

( $\hat{q}$ を標本比率、 $N$ を標本サイズとする)

# $R(a|b)$ の95%信頼区間

コンフュージョン マトリックス	$a / b$	validated $b$	
		植生 (X=1)	非植生 (X=0)
classified $a$	植生 (X=1)	$TP$	$FP$
	非植生 (X=0)	$FN$	$TN$

$$\frac{TN}{FP + TN} - 1.96 \sqrt{\frac{TN}{FP + TN} \left(1 - \frac{TN}{FP + TN}\right) / (FP + TN)} \leq R(0|0) \leq \frac{TN}{FP + TN} + 1.96 \sqrt{\frac{TN}{FP + TN} \left(1 - \frac{TN}{FP + TN}\right) / (FP + TN)}$$

$$\frac{FP}{FP + TN} - 1.96 \sqrt{\frac{FP}{FP + TN} \left(1 - \frac{FP}{FP + TN}\right) / (FP + TN)} \leq R(1|0) \leq \frac{FP}{FP + TN} + 1.96 \sqrt{\frac{FP}{FP + TN} \left(1 - \frac{FP}{FP + TN}\right) / (FP + TN)}$$

$$\frac{FN}{TP + FN} - 1.96 \sqrt{\frac{FN}{TP + FN} \left(1 - \frac{FN}{TP + FN}\right) / (TP + FN)} \leq R(0|1) \leq \frac{FN}{TP + FN} + 1.96 \sqrt{\frac{FN}{TP + FN} \left(1 - \frac{FN}{TP + FN}\right) / (TP + FN)}$$

$$\frac{TP}{TP + FN} - 1.96 \sqrt{\frac{TP}{TP + FN} \left(1 - \frac{TP}{TP + FN}\right) / (TP + FN)} \leq R(1|1) \leq \frac{TP}{TP + FN} + 1.96 \sqrt{\frac{TP}{TP + FN} \left(1 - \frac{TP}{TP + FN}\right) / (TP + FN)}$$

# 分類誤りの影響を考慮した 95%信頼区間推定

2014-2016, 250m

- MGCI(Kapos4) = 0.958986 (0.712942以上)
- MGCI(Kapos5) = 0.994883 (0.745065以上)
- MGCI(Kapos6) = 0.991279 (0.742132以上)

←FAO試算値 0.9968

←FAO試算値 0.9969

←FAO試算値 0.9913

期間	解像度	Kapos分類	標本サイズ	MGCI	MGCI95% 信頼区間 (下限)	MGCI95% 信頼区間 (上限)	R(1 1)	R(1 1)の 95%信頼 区間 (下 限)	R(1 1)の 95%信頼 区間 (上 限)	R(1 0)	R(1 0)の 95%信頼 区間 (下 限)	R(1 0)の 95%信頼 区間 (上 限)	MGCI補正 値	MGCI補正 値の95% 信頼区間 (下限)	MGCI補正 値の95% 信頼区間 (上限)
2014-2016	250m	Kapos1	0	0			0.994278	0.988686	0.998695	0.272727	0.184997	0.360458			
		Kapos2	45	0	0	0	0.994278	0.988686	0.998695	0.272727	0.184997	0.360458	0	0	0
		Kapos3	7163	0.414491	0.403083	0.4259	0.994278	0.988686	0.998695	0.272727	0.184997	0.360458	0.196471	0.033823	0.35912
		Kapos4	172726	0.964684	0.963814	0.965554	0.994278	0.988686	0.998695	0.272727	0.184997	0.360458	0.958986	0.712942	1.20503
		Kapos5	511957	0.990585	0.990321	0.99085	0.994278	0.988686	0.998695	0.272727	0.184997	0.360458	0.994883	0.745065	1.2447
		Kapos6	3438805	0.987985	0.98787	0.9881	0.994278	0.988686	0.998695	0.272727	0.184997	0.360458	0.991279	0.742132	1.240426

# 集計結果 JAXA-LCMGCI (2006-2011)

	2006-2011					
Kapos分類	100m			50m		
	MGCI	MGCI95%下限	MGCI95%上限	MGCI	MGCI95%下限	MGCI95%上限
Kapos1						
Kapos2	0	0	0	0	0	0
Kapos3	0.43661465	0.387174549	0.486054741	0.501464543	0.420215612	0.582713475
Kapos4	0.90404689	0.841974899	0.966118873	0.919162948	0.813461901	1.024863995
Kapos5	0.98893625	0.924181367	1.05369113	0.998846787	0.888282268	1.109411306
Kapos6	0.99070248	0.925975642	1.055429314	1.000842868	0.890196058	1.111489678

# 集計結果 (2006-2011)

	2006-2011					
Kapos分類	250m			100m		
	MGCI	MGCI95%下限	MGCI95%上限	MGCI	MGCI95%下限	MGCI95%上限
Kapos1						
Kapos2	0	0	0	0	0	0
Kapos3	0.346138252	0.230989169	0.461287336	0.467474954	0.279440541	0.655509367
Kapos4	0.916705069	0.762046064	1.071364074	0.905287036	0.664679839	1.145894233
Kapos5	1.014913489	0.852338079	1.177488898	0.987613367	0.736699536	1.238527199
Kapos6	1.018167221	0.855568248	1.180766195	0.990498501	0.739305752	1.241691251

# 集計結果 JAXA-LCMGCI(2014-2016)

Kapos分類	2014-2016					
	250m			100m		
	MGCI	MGCI95%下限	MGCI95%上限	MGCI	MGCI95%下限	MGCI95%上限
Kapos1						
Kapos2	0	0	0	0	0	0
Kapos3	0.19647122	0.03382269	0.359119752	0.324336795	0.219973983	0.428699608
Kapos4	0.95898617	0.712942355	1.205029988	0.964553709	0.814380131	1.114727287
Kapos5	0.99488263	0.745065085	1.244700177	0.99418406	0.84194352	1.146424599
Kapos6	0.99127886	0.742131512	1.240426213	0.991236695	0.839297171	1.143176218

# 集計結果 (2014-2016)

	2014-2016					
Kapos分類	250m			100m		
	MGCI	MGCI95%下限	MGCI95%上限	MGCI	MGCI95%下限	MGCI95%上限
Kapos1						
Kapos2	0	0	0	0	0	0
Kapos3	0.301695247	0.12527167	0.478118825	0.4163877	0.304668677	0.528106722
Kapos4	0.95480915	0.709252616	1.200365684	0.961243392	0.811315279	1.111171504
Kapos5	0.994658075	0.744867218	1.244448932	0.994031799	0.841803903	1.146259696
Kapos6	0.991088511	0.741963741	1.240213281	0.99105755	0.839132891	1.142982209

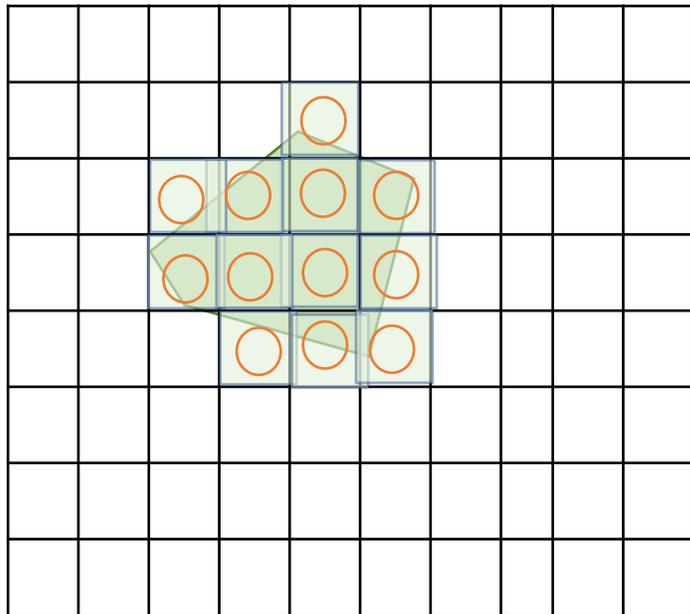
# まとめ

- 2006-2011
  - Kapos4 : 0.905 (0.88以上)
  - Kapos5 : 0.987 (0.88以上)
  - Kapos6 : 0.990 (0.89以上)
- 2014-2016
  - Kapos4 : 0.961 (0.81以上)
  - Kapos5 : 0.994 (0.84以上)
  - Kapos6 : 0.991 (0.84以上)

と思われる。

# 今後の課題

- 筆ポリゴンデータと土地利用被覆分類図グリッドデータとを突合せして分類精度の検証を行う（コンフュージョンマトリックスの標本サイズを増加させる）



- コンフュージョンマトリックスの標本サイズを増加させることにより、 $R(1|1)$ ,  $R(1|0)$ の95%信頼区間幅を狭くする
- その結果、Kapos分類ごとのMGCIの95%信頼区間幅を狭くすることができる