

機械学習を用いた日本域高解像度土地利用・
土地被覆図の高精度化

片木 仁

201310646

平成28年度

筑波大学 生命環境学群生物資源学類

卒業論文

指導教員： 奈佐原顕郎

目 次

第 1 章 序論	1
1.1 はじめに	1
1.2 分類アルゴリズム	2
1.3 解析環境	2
第 2 章 AVNIR-2 画像及び教師点の精査による分類精度向上	4
2.1 背景	4
2.2 目的	4
2.3 衛星画像	4
2.4 手法	6
2.5 検証	14
2.6 考察	14
2.7 結論	16
第 3 章 光学センサと多偏波 SAR を組み合わせた海岸域の砂浜分類	17
3.1 背景	17
3.2 目的	17
3.3 手法	18
3.4 結果	20
3.5 考察	20
第 4 章 多時期の二偏波 SAR を用いた森林・非森林マップの作成	28
4.1 背景	28
4.2 目的	28
4.3 手法	28
4.4 結果	31
4.5 考察	31
第 5 章 結論	42
5.1 まとめ	42
5.2 今後の課題	42
5.2.1 事前確率	42
5.2.2 教師点	43
5.2.3 カテゴリ	43
5.2.4 手法・アルゴリズム	43
5.2.5 検証	44

図 目 次

2.1	AVNIR2HLP に施されている 4 つの補正処理	5
2.2	教師点の土地被覆変化の例。AVNIR2HLP のフォールスカラーで表している。	7
2.3	精査後の教師点の空間分布	8
2.4	山影分布の模式図。山影は時期によって分布が異なる。また山影の箇所は水域と誤分類しやすい。	11
2.5	最終的な土地被覆図	12
2.6	分類全体の流れ	13
2.7	質の悪い教師点を取り除いたことにより改善した土地被覆図。左は取り除く前で、右は取り除いた後。場所は沖縄県島尻郡のチジン山。	14
2.8	震災後の AVNIR2HLP を取り除いたことにより改善した土地被覆図。左は取り除く前で、右は取り除いた後。場所は福島県相馬市の沿岸。	15
3.1	砂浜が都市と誤分類されている土地被覆図の例	18
3.2	AVNIR2HLP の分類に使用した教師点の空間分布	21
3.3	多偏波 SAR の分類に使用した教師点の空間分布	22
3.4	AVNIR2HLP と多偏波 SAR の分類のフローチャート	23
3.5	都市・砂浜の教師点におけるアルファ角の箱ひげ図	24
3.6	高戸前浜海岸の土地被覆図と衛星画像	25
3.7	大洗サンビーチの土地被覆図と衛星画像	26
3.8	日立港周辺の土地被覆図と衛星画像	27
4.1	二偏波 SAR の分類に用いた教師点の空間分布	29
4.2	教師点における HH 偏波の γ_0 の箱ひげ図。 γ_0 は窓サイズ 3 画素 × 3 画素で空間平均を取って計算した。	33
4.3	教師点における HH 偏波の γ_0 の箱ひげ図。 γ_0 は窓サイズ 5 画素 × 5 画素で空間平均を取って計算した。	34
4.4	教師点における HV 偏波の γ_0 の箱ひげ図。 γ_0 は窓サイズ 3 画素 × 3 画素で空間平均を取って計算した。	35
4.5	教師点における HV 偏波の γ_0 の箱ひげ図。 γ_0 は窓サイズ 5 画素 × 5 画素で空間平均を取って計算した。	36
4.6	教師点における γ_0 の箱ひげ図。 γ_0 は時間方向に平均を取って計算した。	37
4.7	多時期二偏波 SAR を用いて分類した FNF (森林・非森林) マップと衛星画像。FNF マップは 10 カテゴリから作成。	38
4.8	多時期二偏波 SAR を用いて分類した FNF (森林・非森林) マップと衛星画像。FNF マップは 3 カテゴリから作成。	39
4.9	多時期二偏波 SAR を用いて分類した FNF (森林・非森林) マップと衛星画像。FNF マップは 3 カテゴリから作成。場所は霞ヶ浦周辺。赤く囲った箇所はハス田。	40

4.10 多時期二偏波 SAR を用いて分類した FNF（森林・非森林）マップと衛星画像。
FNF マップは 3 カテゴリから作成。場所は東海村周辺. 41

表 目 次

1.1 解析に用いたソフトウェアとそのバージョン	3
2.1 AVNIR-2 のセンサ概要	4
2.2 位置ずれのために分類に適さないと判断し、取り除いた AVNIR2HLP の枚数	7
2.3 震災の影響を受けたために分類に適さないと判断し、取り除いた AVNIR2HLP の枚数	9
2.4 分類に使用した AVNIR2HLP の枚数	10
2.5 分類に用いた先駆情報	12
2.6 AVNIR2HLP・教師点の精査前の土地被覆図の混同行列	15
2.7 AVNIR2HLP・教師点の精査後の土地被覆図の混同行列	16
3.1 PALSAR の多偏波モードの概要	18
3.2 分類に使用した多偏波 SAR データの概要	19
3.3 山口四成分分解が想定する散乱過程とその観測対象物	19
4.1 分類に使用した SAR 画像の概要	30

第1章 序論

1.1 はじめに

土地利用・土地被覆図は、土地利用と土地被覆を記述した地図である。土地被覆は「地球上で観測される（生物）物理的被覆」と定義されている[1]。また土地利用は土地がどのように利用されているかを表す。以下では土地利用・土地被覆をまとめて土地被覆と扱い、その地図を土地被覆図とする。

土地被覆図は様々な分野で使われている。例えば農業利用や生態系評価、防災や環境シミュレーションだ。土地被覆図はその利用目的によって、記述した内容（土地被覆のカテゴリ）が異なる。例えば農業利用では、田畠の種類や耕作放棄地が書かれている土地被覆図が用いられる。また地図の細かさ（空間解像度）も数 m～数十 m と細かい方が役立つ。一方で地球規模の環境シミュレーションでは、数百 m 以上の荒い土地被覆図が用いられる。

土地被覆図は衛星画像を用いて作られることが多い。これは衛星画像が (i) 広範囲をカバーし (ii) 定期的に得られるからである。しかしながら衛星画像を用いた地球全体の土地被覆図の精度は 70%前後であり [4], 十分ではない。

日本でも宇宙航空分野の研究開発を行っている JAXA をはじめとするいくつかの機関が土地被覆図を作成している。JAXA では機械学習の手法である教師あり分類を用いている。ここにおける機械学習とは、人間が教えたルールに基づいて機械が土地被覆分類を行うことである。教師あり分類は教師点（ある地点の土地被覆が何であるかを人間が判断し、機械に教える空間上の点）に基づいて機械が学習する手法である。

JAXA では衛星画像の反射率に基づいた分類に加え、ベイズ推定に基づいた先駆情報を取り入れて精度の向上を試みた [2]。先駆情報とは人間が経験的に得られる情報（例えば河川の脇には草地が多い、夜間に明るい箇所は都市など）である。

ベイズ推定は以下の式で表される [3],

$$p(C_k|\mathbf{x}) = \frac{p(\mathbf{x}|C_k)p(C_k)}{\sum p(\mathbf{x})} \quad (1.1)$$

ここで \mathbf{x} : 入力情報, C_k : 土地被覆カテゴリであり, $p(C_k|\mathbf{x})$: 事後確率, $p(\mathbf{x}|C_k)$: 尤度, $p(C_k)$: 事前確率である。先駆情報を $p(C_k)$ に反映させることにより分類精度は向上してきたが、未だ改善点が多い。

さて、土地被覆図の作成には種々の要因がからみ合っている。例えば入力画像の質（位置ずれ・大気補正・地形補正）や教師データ（総数・カテゴリ間の数の比率・空間代表性），分類手法と，どの要因が土地被覆図の精度に影響を及ぼしているかを一度に考慮するのは難しい。

そこで本研究では土地被覆図の精度を向上させるために、多面的な検討を行った。まずはじめに入力に用いた教師点・衛星画像の精査を行った（第2章）。続いて海岸域の砂地に着目し、誤分類の低減を行った。（第3章）。また多時期SAR画像の分類手法を検討するため、森林・非森林マップを作成した（第4章）。最後に以上の結果をまとめ、今後の課題を述べた（第5章）。

1.2 分類アルゴリズム

衛星画像の分類には、教師あり分類の手法であるカーネル密度推定[3]（以下、KDE）を用いた。

$$p(\mathbf{x}|C_k) = \frac{1}{N_{C_k}} \sum_{n=1}^{N_{C_k}} \left\{ \prod_{d=1}^D \frac{1}{h_d} K \left(\frac{x_d - y_{n,d}}{h_d} \right) \right\} \quad (1.2)$$

ただし $p(\mathbf{x}|C_k)$: カテゴリ C_k の尤度、 \mathbf{x} : 入力画像の各ピクセルにおける特徴ベクトル、 N_{C_k} : カテゴリ C_k の教師点の総数、 D : 特徴空間の次元数、 x_d : 入力画像の各ピクセルにおける特徴量、 $y_{n,d}$: 教師データの特徴量である。特徴は入力画像からカテゴリを分類するために必要な情報であり、それを数値化したものが特徴量である。特徴量をまとめて並べたものを特徴ベクトルと呼ぶ。また特徴空間は、特徴ベクトルにより貼られる空間である。

カーネルはガウス関数を用いた、

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} u^2 \right) \quad (1.3)$$

またガウス関数のバンド幅 h_d は以下のスコットの経験則により求めた[3]、

$$h_d = N^{\frac{-1}{D+4}} \sigma_d \quad (1.4)$$

ここで σ_d : 教師データにおける次元 d の標準偏差であり、 N : 全カテゴリの教師点の総数である。

衛星画像はデータサイズが大きく、KDE分類は時間がかかる。そこでJAXAの土地被覆分類ソフトウェアであるsaclassを用いた。saclassはGPGPU (General-purpose computing on graphics processing units; GPUによる汎用計算)を利用した。それにあたり、NVIDIA社のCUDAという統合開発環境を使用した。これにより高速に衛星画像を分類することができる。

1.3 解析環境

本研究で使用したOS・ソフトウェアのバージョンを表1.1に示した。なお第2章と第3章・第4章でバージョンが異なるのは、便宜上、異なる計算機の環境で解析を行ったためである。

表 1.1: 解析に用いたソフトウェアとそのバージョン

	第 2 章	第 3 章・第 4 章
Linux OS	CentOS 6.6	Xubuntu 16.04
saclass	1.2	1.3
GRASS GIS	6.4.4	7.0.3
QGIS	1.8.0	2.18.1
GDAL	1.11.1	1.11.3
SQLite3	3.6.20	3.11.0
Python	-	3.5.2
matplotlib	-	1.5.3
numpy	-	1.11.2
pandas	-	0.19.1
seaborn	-	0.7.1

第2章 AVNIR-2画像及び教師点の精査による分類精度向上

2.1 背景

JAXA の日本域土地被覆図は、分類に用いてきた教師点・衛星画像のチェックが不十分であった。質の悪い教師点・衛星画像が混ざっていると、誤ったカテゴリで土地被覆図を作成してしまう可能性がある。

2.2 目的

教師点・衛星画像の質を精査することにより、分類精度の向上を試みることを目的とした。

2.3 衛星画像

ALOS/AVNIR-2の観測画像に補正を施したAVNIR2 High-Level Products（以下、AVNIR2HLP）を分類に用いた。AVNIR-2はJAXAの光学センサであり（表2.1）、1ピクセルの大きさを表す空間解像度は約10mである。画像に施されている補正是ラジオメトトリック補正・オルソ幾何補正・大気補正・斜面補正である（図2.1）。ラジオメトトリック補正是衛星が受信した電気信号を、物理量（放射輝度）に直す処理である。オルソ幾何補正是地形の起伏や地球の表面が曲面であることによって生じる歪みを直し、地図に投影する。大気補正是大気粒子による散乱の影響を補正し、斜面補正是斜面と太陽の向きによって変化する日陰及び日向を補正する。これらの処理を施すことにより、分類精度が向上することが期待される。

表 2.1: AVNIR-2 のセンサ概要

バンド番号	波長 ($\mu\text{ m}$)
Band1	0.42 ~ 0.50
Band2	0.52 ~ 0.60
Band3	0.61 ~ 0.69
Band4	0.76 ~ 0.89

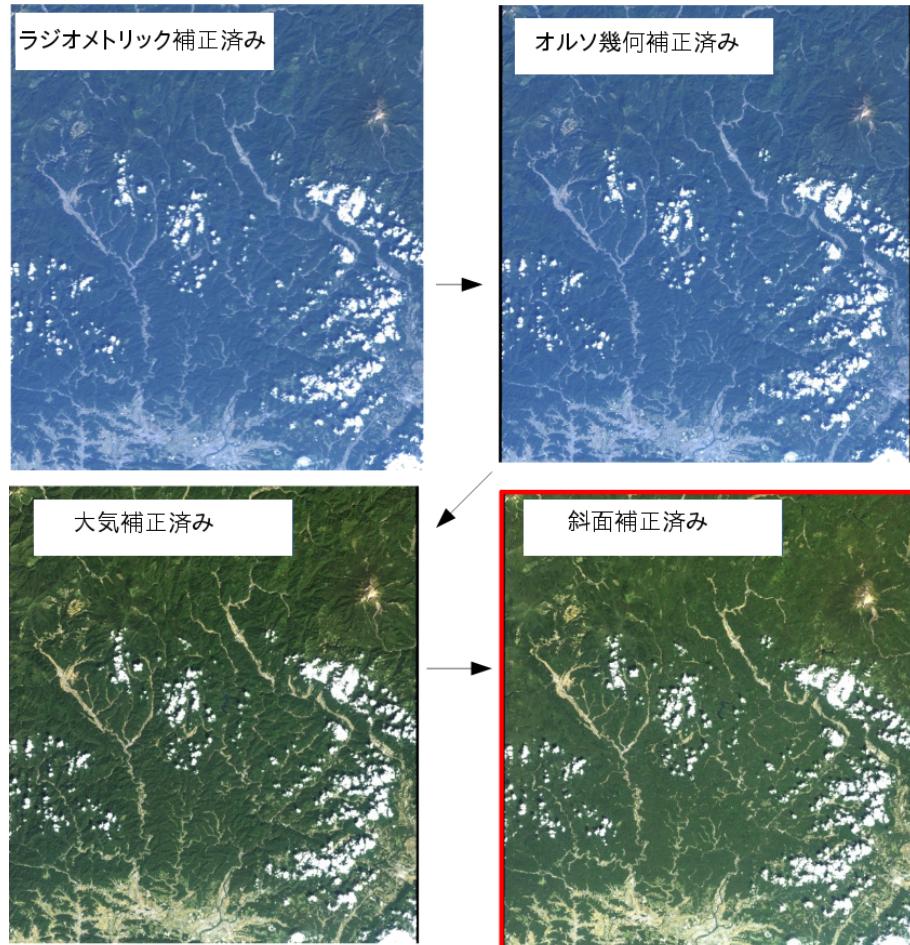


図 2.1: AVNIR2HLP に施されている 4 つの補正処理

2.4 手法

AVNIR2HLP を日本全域で 1 度 × 1 度の計 126 タイルに分割して分類した。

AVNIR-2 の観測期間は 2006 年の 5 月から 2011 年の 4 月末までの 5 年間である。本研究では AVNIR-2 の全観測期間の AVNIR2HLP を用い、2006 年から 2011 年までの代表的な様子を表す一枚の土地被覆図を作成した。各年の土地被覆図を作成するのに比べて分類時の情報が増え、その結果、分類精度が向上するためである。

作成の際に重要なのはカテゴリ分けである。これが細かすぎると、土地被覆図の作成に労力がかかる。かと言って大まかにしすぎると、利用目的が限られてしまう。そこで今回はその間を取り、10 種類とした。具体的には水域 (Water)・都市 (Urban)・水田 (Paddy)・畑地 (Crop)・草地 (Grass)・落葉広葉樹 (DBF)・落葉針葉樹 (DNF)・常緑広葉樹 (EBF)・常緑針葉樹 (ENF)・裸地 (Bareland) である。

はじめに緯度・経度・カテゴリの 3 つを含む教師点を作成した。

教師点は SACL AJ 及び Google Satellite 画像の判読から取得した。SACL AJ は現地に赴いて写真を撮影し、その土地被覆カテゴリが何かを撮影者が判読し登録する土地被覆検証データサイトである。このような現地踏破情報によるカテゴリは、衛星画像を判読しカテゴリを取得するより確実である。しかしながら SACL AJ のみでは (i) 教師点の数が十分に得られず (ii) 空間的な偏りが生じてしまうため、補助的に Google Satellite 画像を教師点の取得に用いた。Google Satellite 画像は商用衛星や航空写真で撮影された画像である。そのため解像度が数十 cm と非常に細かく、土地被覆のカテゴリが何かを判読しやすい。SACL AJ と Google Satellite 画像を合わせて、教師点を合計 20,332 点取得した。

続いて取得した教師点を精査した。具体的には AVNIR2HLP の 2006 年の 5 月から 2011 年の 4 月末の画像から、教師点の周辺を切り取って小画像を作成した。作成した小画像を動画にし、教師点が土地被覆変化をしていないか・カテゴリの判別を間違えていないか・教師点同士が隣接していないか（基準：50 m）を 20,332 点全てを目視で確認した。土地被覆が変化したりカテゴリの判別を間違えた教師点が混ざっていると、異なるカテゴリの特徴量を元に分類してしまう。また教師点同士があまりに近いと空間分布の偏りが生じてしまい、誤分類の原因になる可能性がある。

目視の結果、土地被覆が変化する教師点が見つかった（図 2.2）。例えば茨城県つくば市周辺に 2006 年の時点では草原であったが、2010 年には都市になった教師点があった。また沖縄県で 2007 年の時点ではダムが完成していなかったが、2010 年にはダムが完成し水域となつた教師点が存在した。目視により合計 1,696 点の教師点を分類に適さないと判断し除外し、18,636 点を分類に用いた（図 2.3）。

さらに教師点の精査により、位置ずれが大きい AVNIR2HLP が計 25 枚見つかったので、除外した（表 2.2）。ここでタイル名は画像の範囲を表しており、N36E140 であれば北緯 36 度東経 140 度から北緯 37 度東経 141 度までの 1 度 × 1 度の範囲を表す。位置ずれが大きい AVNIR2HLP を取り除いた理由は (i) 教師データの作成時に、異なるカテゴリの特徴量を取得する恐れがあり (ii) 入力画像の分類結果の統合時に、位置の異なる分類結果を掛けあわせてしまい、誤分類につながるためである。

また 2011 年の 3 月に東日本大震災が発生した。この震災により東北地方の沿岸部で津波

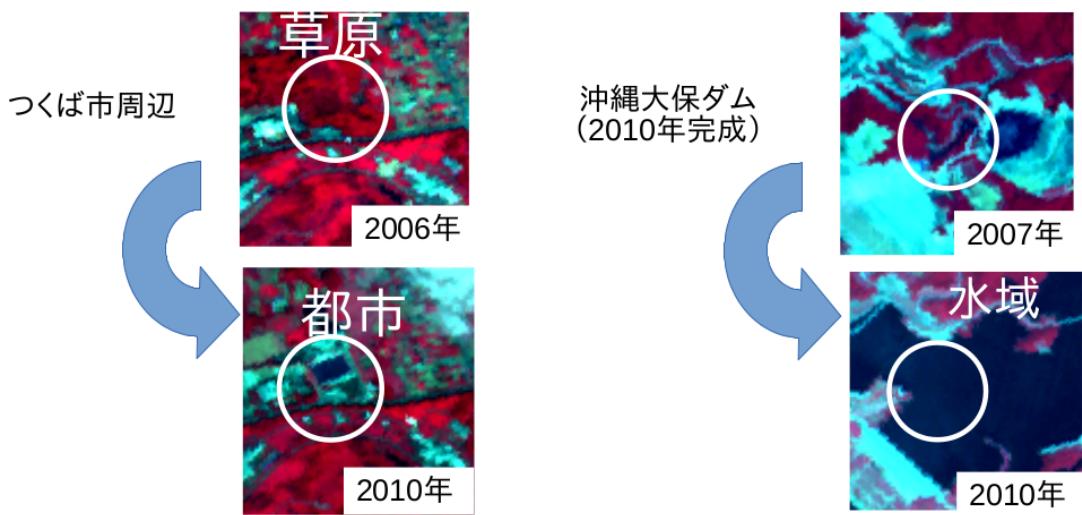


図 2.2: 教師点の土地被覆変化の例. AVNIR2HLP のフォールスカラーで表している.

表 2.2: 位置ずれのために分類に適さないと判断し, 取り除いた AVNIR2HLP の枚数

タイル名	取り除いた枚数
N24E141	1
N27E142	1
N33E132	1
N33E136	2
N34E135	1
N34E136	4
N34E137	1
N35E133	1
N35E139	1
N35E140	1
N36E136	1
N36E140	1
N39E139	2
N39E140	5

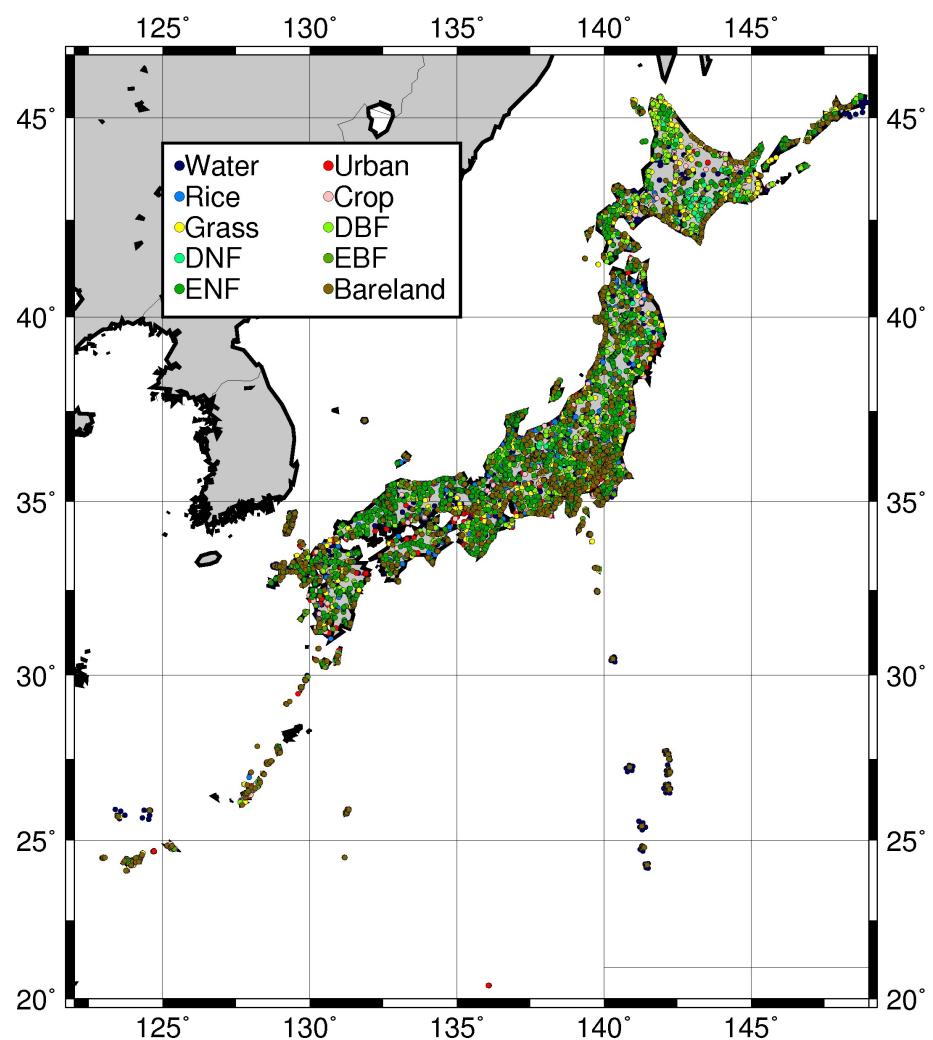


図 2.3: 精査後の教師点の空間分布

表 2.3: 震災の影響を受けたために分類に適さないと判断し、取り除いた AVNIR2HLP の枚数

タイル名	取り除いた枚数
N37E140	30
N37E141	25
N38E141	25
N39E141	16
N39E142	17
N40E141	9

の被害が起き、その結果沿岸部の土地被覆は大きく変化した。繰り返しになるが、AVNIR-2 の観測期間は 2006 年 5 月から 2011 年の 4 月末までである。観測開始から震災が起こる前までの期間と、震災後から観測終了までの期間の長さを比べると、起こる前の方が長い。したがって、2006 年から 2011 年までの代表的な様子を表すためには、震災前までの AVNIR2HLP を用いた方がよいと考えた。そこで東北地方の震災後の AVNIR2HLP を計 122 枚取り除いた（表 2.3）。

最終的に分類に使用した AVNIR2HLP の枚数を表 2.4 に示した。

続いて教師点を元に、分類に必要な教師データを AVNIR2HLP を参照して作成した。教師データは、カテゴリ・BAND1 から BAND4 の反射率・衛星データの観測日を格納した。

そして教師データを元に KDE 分類（第 1 章参照）を行い、AVNIR2HLP の各シーンに対してカテゴリの尤度を計算した。特徴空間は {BAND1 の反射率, BAND2 の反射率, BAND3 の反射率, BAND4 の反射率, t_1, t_2 } の 6 次元である。ただし t_1, t_2 は AVNIR2HLP の観測日の年間通算日（day of year; 以下 DOY）を二次元の単位円上で表したものである [3]。

$$[t_1, t_2] = \left[\cos\left(2\pi \frac{\text{DOY}}{\text{DOY}_{\text{total}}}\right), \sin\left(2\pi \frac{\text{DOY}}{\text{DOY}_{\text{total}}}\right) \right] \quad (2.1)$$

ただし DOY_{total} は 12 月 31 における年間通算日であり、うるう年であれば 366、そうでない年は 365 である。観測日を二次元の単位円上で表すことにより、年によらず周期的に同じような地表面の様子を表すという季節性を分類の特徴量として取り入れることができる。入力値 x_d は入力画像の BAND1 から BAND4 の反射率, t_1, t_2 である。また分類に用いた教師データの使用範囲は 3 度 × 3 度である。例えば N36E140 のタイルの教師データの使用範囲は、北緯 35 度東経 139 度から北緯 38 度東経 142 度に含まれるものである。

そして各シーンの尤度に対し、山影処理を施した。山影により暗くなる箇所は、時期によって異なる。また山影と日向は一般に異なる反射率を示す。そのため、山影の箇所は誤分類されがちである（図 2.4）。そこで山影のシーンの重要度を下げ、日向のシーンの重要度を上げる処理を取り入れた [2]。具体的には数値地形データ（DEM）と AVNIR2HLP の DOY によって山影を推定し、以下の線形変換を行った。

$$p_{\text{after}} = a \times p(\mathbf{x}|C_k) + \frac{(1-a)}{M} \quad (2.2)$$

表 2.4: 分類に使用した AVNIR2HLP の枚数

タイル名	使用した枚数	タイル名	使用した枚数	タイル名	使用した枚数
N20E136	14	N33E130	128	N37E141	64
N24E122	4	N33E131	125	N38E138	75
N24E123	25	N33E132	99	N38E139	101
N24E124	42	N33E133	100	N38E140	127
N24E125	19	N33E134	86	N38E141	76
N24E131	18	N33E135	81	N39E139	64
N24E141	31	N33E136	70	N39E140	112
N24E153	5	N33E139	53	N39E141	104
N25E123	16	N34E129	61	N39E142	41
N25E124	40	N34E130	83	N40E139	39
N25E131	5	N34E131	86	N40E140	95
N25E141	32	N34E132	87	N40E141	92
N26E126	17	N34E133	106	N41E139	52
N26E127	74	N34E134	108	N41E140	98
N26E128	56	N34E135	122	N41E141	71
N26E142	26	N34E136	131	N41E143	16
N27E127	43	N34E137	111	N42E139	61
N27E128	63	N34E138	91	N42E140	99
N27E129	40	N34E139	100	N42E141	88
N27E140	3	N35E132	57	N42E142	71
N27E142	23	N35E133	87	N42E143	73
N28E128	50	N35E134	80	N42E144	56
N28E129	66	N35E135	102	N42E145	40
N28E130	23	N35E136	123	N43E140	68
N29E129	61	N35E137	139	N43E141	101
N30E129	60	N35E138	142	N43E142	102
N30E130	85	N35E139	159	N43E143	108
N30E131	43	N35E140	126	N43E144	109
N30E140	20	N36E132	33	N43E145	92
N31E128	39	N36E133	38	N43E146	71
N31E129	86	N36E135	52	N44E141	100
N31E130	129	N36E136	91	N44E142	113
N31E131	94	N36E137	126	N44E143	74
N32E128	68	N36E138	140	N44E144	52
N32E129	112	N36E139	146	N44E145	68
N32E130	147	N36E140	164	N44E146	96
N32E131	121	N37E131	10	N44E147	59
N32E132	65	N37E136	54	N45E140	29
N32E133	44	N37E137	88	N45E141	75
N32E139	40	N37E138	115	N45E142	72
N33E128	44	N37E139	117	N45E147	46
N33E129	97	N37E140	108	N45E148	40

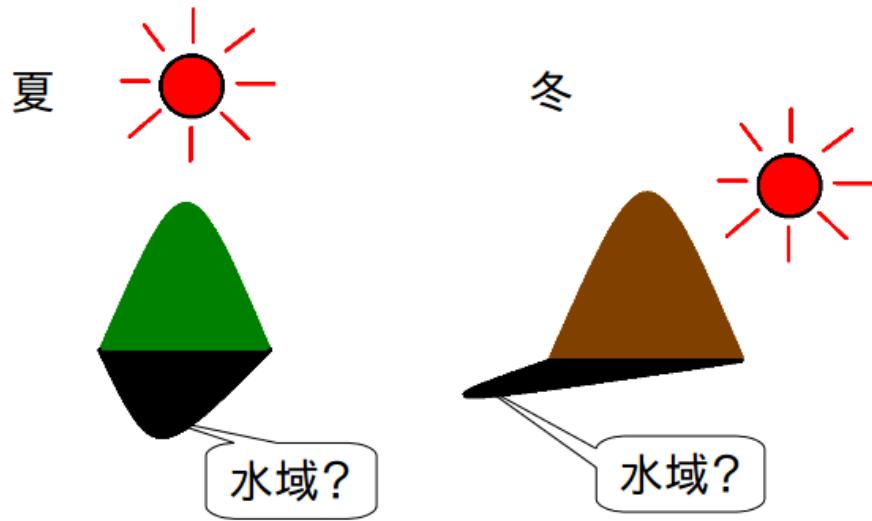


図 2.4: 山影分布の模式図. 山影は時期によって分布が異なる. また山影の箇所は水域と誤分類しやすい.

ただし,

$$a = \begin{cases} 0.7 & (\text{日向の場合}) \\ 0.1 & (\text{山影の場合}) \end{cases} \quad (2.3)$$

であり, $p(\mathbf{x}|C_k)$: 山影処理前の尤度 (式 1.2), p_{after} : 山影処理後の尤度である. また M はカテゴリの数で, ここでは 10 とした.

重み変数 a は誤分類が少なくなるように試行を繰り返して求めた. また DEM は PRISM DSM の 10 m 解像度を用いた.

その後, 山影処理を施した尤度をシーン同士で掛けあわせ, 一枚の尤度マップを作成した.

次に表 2.5 に示した先駆情報に対して KDE 分類を行い, 先駆情報で一枚の事前確率マップを作成した. ただし鉄道網は鉄道のみを使用するためトンネルを, 水路網は河川のみを使用するために用水路を除外した.

そして尤度マップと事前確率マップを掛けあわせて事後確率マップを作成した. その後 3 画素 × 3 画素の多数決フィルタによりランダムノイズを取り除き, 事後確率値が最も高いカテゴリを分類結果とした. 最後に分類結果と Google Satellite 画像を見比べ, 砂浜が都市と分類されているような明らかな誤分類を目視にて判読し, 正しいカテゴリになるように手で編集した. そして手で編集した結果を最終的な土地被覆図とした (図 2.5).

以上の処理の全体を, 図 2.6 に示した.

表 2.5: 分類に用いた先駆情報

分類に用いた先駆情報	いつのデータ	空間解像度	オリジナルのデータ
道路網の片道 2500 m の距離バッファ	2016 年 02 月 02 日 20:13:01Z	10 m	OpenStreetMap [5]
鉄道網の片道 2500 m の距離バッファ	2016 年 02 月 02 日 20:13:01Z	10 m	OpenStreetMap [5]
水路網の片道 2500 m の距離バッファ	2016 年 02 月 02 日 20:13:01Z	10 m	OpenStreetMap [5]
建物の片道 2500 m の距離バッファ	2016 年 02 月 02 日 20:13:01Z	10 m	OpenStreetMap [5]
北海道市町村ごとの水稻作付の有無情報	2014 年	20 m	農林水産関係市町村別統計 [6]
HV 偏波モザイク	2008 年	25 m	ALOS/PALSAR
海岸線情報	-	10 m	ALOS/PRISM
傾斜のラスター	-	10 m	国土地理院数値地形データ
夜間光データ	2015 年 10 月 09 日 - 12 月 20 日	500 m	Suomi NPP Night-light
先駆マップ	-	20 m	その土地に詳しい判読者の情報

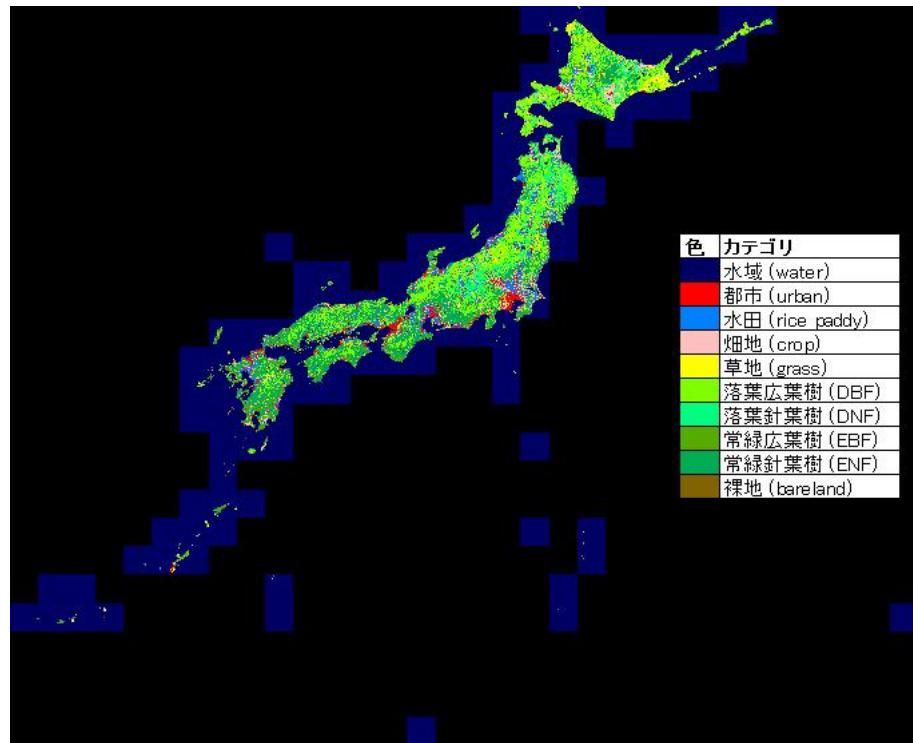


図 2.5: 最終的な土地被覆図

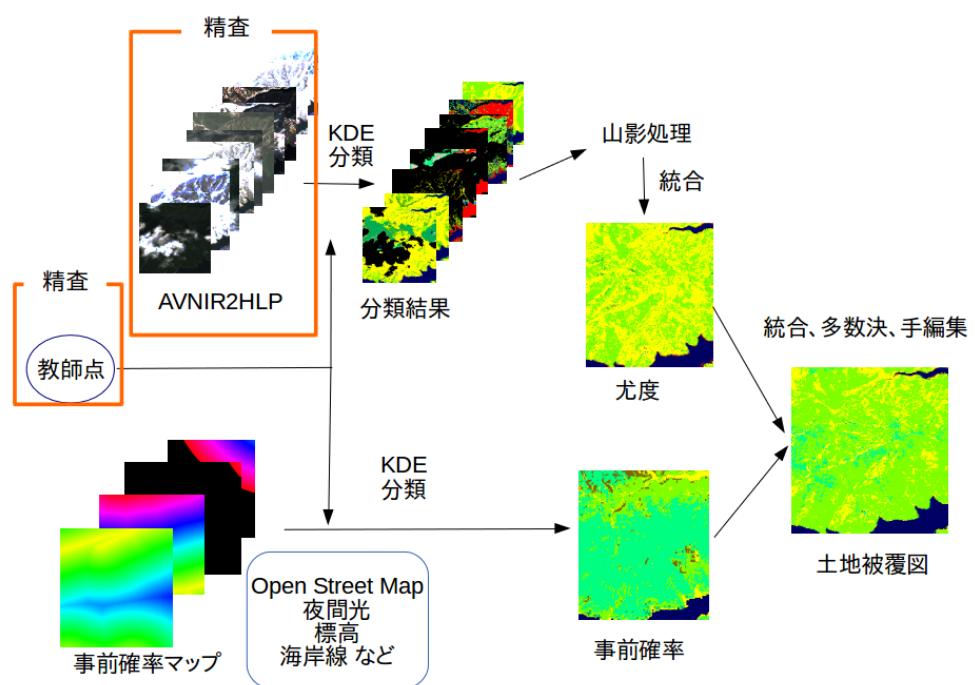


図 2.6: 分類全体の流れ

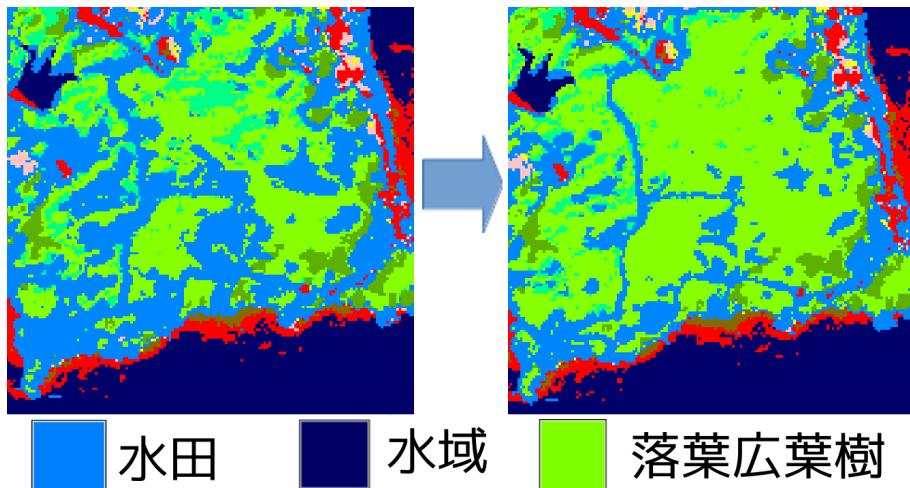


図 2.7: 質の悪い教師点を取り除いたことにより改善した土地被覆図. 左は取り除く前で, 右は取り除いた後. 場所は沖縄県島尻郡のチジン山.

2.5 検証

まずははじめに目視によって分類結果を検証した. 質の悪い教師点を取り除く前後で土地被覆図がどう変わったかを示した(図 2.7). 左の図では森林域に偽の水田が入りこんでいた. しかし質の悪い教師点を取り除いた後の右の図ではその偽の水田を取り除くことができた.

続いて震災後の AVNIR2HLP を取り除く前後で土地被覆図がどう変わったかを示した(図 2.8). 左の図では震災の影響で水田が水域に, 常緑広葉樹が都市・水田と分類されていた. しかし震災後の AVNIR2HLP を取り除いたことにより, 震災前までの代表的な状況を表すことができた.

続いて混同行列を用いて精度検証を行った. 検証点は SACLAD 等から教師点とは独立に 1409 点を作成した. 入力データを精査する前の全体精度は 76.9% で, 分類の一一致率の指標である κ 係数は 0.733 であった(表 2.6). しかし精査後は全体精度は 78.0%, κ 係数は 0.745 と改善した(表 2.7).

2.6 考察

教師点を精査した結果, Google Satellite 画像判読・SACLAD から取得した 20,332 点のうち, 全体の 8% (1696 点) が分類に適さないと判断された. Google Satellite 画像は複数時期の画像をモザイク状に貼り付けた一枚の画像である.そのため, 特定の地域の画像を多時期で参照することはできない. さらに画像の時期も AVNIR-2 の観測期間と一致していない

¹PA : Producer's Accuracy

²UA : User's Accuracy

³OA : Overall Accuracy

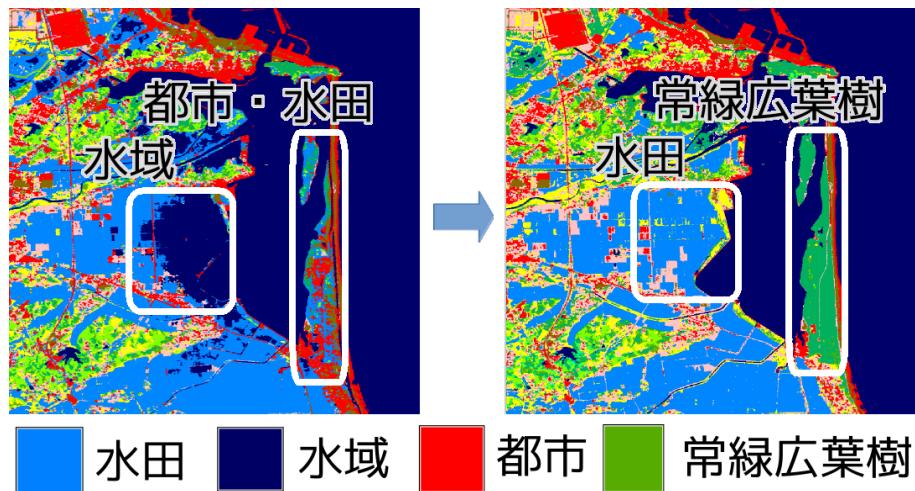


図 2.8: 震災後の AVNIR2HLP を取り除いたことにより改善した土地被覆図. 左は取り除く前で、右は取り除いた後. 場所は福島県相馬市の沿岸.

表 2.6: AVNIR2HLP・教師点の精査前の土地被覆図の混同行列

	Results											PA ¹ (%)	
	Water	Urban	Paddy	Crop	Grass	DBF	DNF	EBF	ENF	Bareland	Total		
Reference	Water	193	2	1	0	0	0	0	0	1	197	98.0	
	Urban	2	221	2	1	0	0	0	0	3	229	96.5	
	Paddy	1	2	260	18	8	1	0	0	1	291	89.3	
	Crop	2	2	38	71	39	5	0	6	0	168	42.3	
	Grass	0	0	8	15	43	5	2	7	0	81	53.1	
	DBF	0	1	2	6	8	65	19	17	13	0	131	49.6
	DNF	0	0	0	0	1	1	11	0	1	0	14	78.6
	EBF	0	0	0	0	1	4	1	33	16	0	55	60.0
	ENF	1	0	0	1	0	12	5	29	166	0	214	77.6
	Bareland	0	4	1	1	2	0	0	0	21	29	72.4	
Total		199	232	312	113	102	93	38	92	196	32	1409	-
UA ² (%)		97.0	95.3	83.3	62.8	42.2	69.9	28.9	35.9	84.7	65.6	-	OA ³ : 76.9 (%)

表 2.7: AVNIR2HLP・教師点の精査後の土地被覆図の混同行列

	Results												
	Water	Urban	Paddy	Crop	Grass	DBF	DNF	EBF	ENF	Bareland	Total	PA ¹ (%)	
Reference	Water	193	1	1	0	0	0	0	0	2	197	98.0	
	Urban	2	222	2	1	0	0	0	0	2	229	96.9	
	Paddy	1	2	260	18	6	2	0	1	0	291	89.3	
	Crop	1	2	28	76	41	4	1	9	1	168	45.2	
	Grass	0	0	10	14	42	4	2	8	0	81	51.9	
	DBF	0	1	2	8	5	74	13	15	13	131	56.5	
	DNF	0	0	0	0	1	1	11	0	1	0	14	78.6
	EBF	0	0	0	0	1	5	1	32	16	0	55	58.2
	ENF	1	0	0	1	0	11	4	30	167	0	214	78.0
	Bareland	0	2	1	1	3	0	0	0	22	29	75.9	
Total		198	230	304	119	99	101	32	95	198	33	1409	-
UA ² (%)		97.5	96.5	85.5	63.9	42.4	73.3	34.4	33.7	84.3	66.7	-	OA ³ : 78.0 (%)

可能性がある。また SACL AJ は多時期の現地踏破情報を整備することを目的としているが、本研究では単時期の情報しか利用していなかった。そのため土地被覆が変化する教師点が混ざっていたと考えられる。

また教師点・入力画像の精査を行ったことにより、全体精度・ κ 係数ともに良くなつた。このことから教師点・入力画像が分類精度の向上に必要だと考えられる。

2.7 結論

土地被覆図の精度向上のためには、教師点の土地被覆変化がないかを確認する必要がある。

第3章 光学センサと多偏波SARを組み合わせた海岸域の砂浜分類

3.1 背景

JAXA の日本域土地被覆図は、光学センサで観測された画像 (AVNIR2HLP) を元に分類を行ってきた。しかし波打ち際の砂浜は、しばしば都市と誤分類されてしまう（図 3.1）。これは波打ち際の砂浜は白く泡立ち、その光学的特性が都市と似ているためである。

合成開口レーダー (Synthetic Aperture Radar; SAR) は、アンテナから発するマイクロ波を観測対象物に当て、その反射を受信する能動型センサである。合成開口の名の通り、レンジ圧縮とアジャマス圧縮という技術を用いて受信電波を合成することで、受信画像を高解像度化している。SAR の特徴として、昼夜関係なく観測できること、また大気中の分子を透過するため天候に左右されず観測できる点が挙げられる。

さて、マイクロ波を含めて一般的に電磁波は水平 (Horizontal) 偏波と垂直 (Vertical) 偏波の 2 つの独立な偏波の線型結合で状態を表せる。アンテナで水平偏波・垂直偏波を送信・受信することを考えると、送信で水平・垂直の 2 通り、受信で水平・垂直の 2 通りで偏波の組み合わせは合計 4 通りある。4 つの偏波を組み合わせた SAR を多偏波 SAR と呼ぶ。

多偏波 SAR の送信電界と受信電界の関係は、以下の関係式で表せれる。

$$\mathbf{E}^s = [S]\mathbf{E}^t = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} \mathbf{E}^t \quad (3.1)$$

ここで \mathbf{E}^t はアンテナが送信するマイクロ波の電界ベクトル、 \mathbf{E}^s は受信する電界ベクトル、 $[S]$ は Sinclair 散乱行列である。電磁波は横波であり進行方向の成分は 0 となるので、 \mathbf{E}^t 、 \mathbf{E}^s はそれぞれ 2 次元のベクトルとなる。 $[S]$ は水平偏波・垂直偏波を基底とした時の表現行列である。例えば S_{HV} は水平偏波で受信する電界ベクトルの大きさを、垂直偏波で送信した大きさで割った成分である。4 つの偏波情報を用いることにより、観測対象物の散乱過程を記述でき、土地被覆カテゴリを区別できると期待される。

3.2 目的

光学センサで都市と誤分類されてしまう砂浜を、多偏波 SAR を用いて裸地と分類することを目的とした。



(a) AVNIR2HLP で作成した土地被覆図
(b) Google Satellite 画像 (地図データ:Google, DigitalGlobe)

図 3.1: 砂浜が都市と誤分類されている土地被覆図の例

3.3 手法

日本域土地被覆図の作成は、日本全域を対象としている。しかし日本全域で砂浜分類を試すには時間がかかるため、まずは小さな領域で検討する必要がある。そこで北緯 36 度東経 140 度から北緯 37 度東経 141 度の 1 度 × 1 度で、砂浜分類を行った。この地域を選定した理由は、土地被覆図作成の重要なテストサイトとして情報が集約されているからである。

光学センサのデータは AVNIR2HLP を、SAR のデータは ALOS/PALSAR の多偏波モードで観測した画像を土地被覆図の作成に用いた (表 3.1, 3.2),

表 3.1: PALSAR の多偏波モードの概要

項目	値
中間周波数	1270 MHz (L-band)
バンド幅	14 MHz
偏波	HH・HV・VH・VV
入射角範囲	8-30°
地上分解能	24-89 m
観測幅	20-65 km

多偏波 SAR には以下の 3 つの表現がある。この表現の違いにより砂浜がどう分類されるかを調べるために、それぞれで分類を試し比較した。

1. HH, HV, VV の各強度

表 3.2: 分類に使用した多偏波 SAR データの概要

項目	値
観測期間	2009/03/12 - 2009/06/11
オフナディア角	21.5° と 23.1°
軌道	アセンディング
シーン数	564 枚

2. 山口四成分分解に基づく山口四成分

3. Cloude-Pottier 分解に基づくアルファ角

HH, HV, VV の各強度は、各偏波の後方散乱係数 (σ_0 値) である。多偏波 SAR は HH, HV, VH, VV の 4 つの強度を得るが、一般的に HV 強度と VH 強度は等しいため、多偏波 SAR 解析では VH 強度を用いない。そこで本研究でも HH, HV, VV の 3 つの強度を用いた。

山口四成分分解は 4 つの散乱過程（表面散乱・二回散乱・体積散乱・ヘリックス散乱）を表すモデル行列を仮定し、多偏波 SAR によって得られる散乱行列をモデル行列の線型結合として展開する手法である [7]。ここで各モデル行列の係数を、山口四成分とする。各散乱過程とその代表的な観測対象物の対応を表 3.3 に示した [7]。本研究では砂浜は表面散乱が、都市は二回散乱が他の散乱過程より卓越するという仮定のもと、山口四成分分解データを全て用いた。

Cloude-Pottier 分解 [8] は散乱行列の固有値計算に基づき、エントロピー・アルファ角の 2 つの指標を得る手法である。エントロピーは散乱のランダム性を表し、アルファ角は偏波依存性を表す [7]。アルファ角は砂浜では小さく、都市では大きくなると予想される。エントロピーは予備実験で、砂浜と都市を分けるのに有用な特徴量ではないと判明したので、今回はアルファ角のみを分類に用いた。

本研究で使用した各強度のデータはアジマス方向に 4 ルック処理を施し、山口四成分・アルファ角はアジマス方向に 4 ルックした散乱行列を用意し、その散乱行列に 5 画素 x 5 画素のガウシアンフィルタをかけ、そこから各値を算出した (JAXA 処理)。山口四成分・アルファ角でガウシアンフィルタを施した理由は、平均数が少ないとバイアスを除去できないためである [9]。

ここまで多偏波 SAR を用いると都市と砂浜は分離できると仮定したが、多偏波 SAR のみを用いると砂浜と海を誤分類してしまう。これは砂浜と海の散乱過程が似ているためである。

表 3.3: 山口四成分分解が想定する散乱過程とその観測対象物

散乱過程	代表的な観測対象物
表面散乱	海面・農地
二回散乱	建物・樹幹
体積散乱	ランダムに傾いたワイヤーの集合
ヘリックス散乱	人工物

る。一方で光学センサは光学特性の違いから、砂浜と海を分離することができる。そこで第一段階として、まずは光学センサで分類を行った。

具体的にはAVNIR2HLPを第2章と同じ手法(KDE分類)・教師点・カテゴリで分類した。そして都市・裸地以外のカテゴリと分類された箇所では、その結果を土地被覆図とした。以下第2章の繰り返しになるが、カテゴリは水域(Water)・都市(Urban)・水田(Paddy)・畑地(Crop)・草地(Grass)・落葉広葉樹(DBF)・落葉針葉樹(DNF)・常緑広葉樹(EBF)・常緑針葉樹(ENF)・裸地(Bareland)とした。特徴空間はBAND1からBAND4の反射率、 t_1 、 t_2 の6次元とした。ただし t_1 、 t_2 はAVNIR2HLPの観測日のDOYを二次元の単位円上で表したものである(第2章参照)。分類に用いた教師点は第2章で精査したものを使い、教師データの使用範囲は3度×3度、すなわち北緯35度東経139度から北緯38度東経142度に含まれるものとした(図3.2)。

一方AVNIR2HLPが都市または裸地と分類した箇所では、それを無視し、代わりに多偏波SARを分類した。そして多偏波SARの分類結果を参照し、その結果を土地被覆図とした。具体的には都市・裸地の判定に用いるため、カテゴリを都市・裸地の2種類とし、多偏波SARをKDE分類した。特徴空間は(1)HH、HV、VV強度の3次元、(2)山口四成分の4次元、(3)アルファ角の1次元とした。裸地の教師点はAVNIR2HLPで都市として分類された砂浜を選び、Google Satellite画像を参照して緯度・経度を取得し作成した。また、都市の教師点はAVNIR2HLPの分類で用いたもののうち、北緯36度東経140度から北緯37度東経141度に含まれるものを使用した(図3.3)。

以上の処理を図3.4に示した。

3.4 結果

3つのSAR表現で土地被覆図の様子が異なった砂浜に着目し、目視にて確認した(図3.6, 3.7, 3.8)。

(1)および(2)は、同様の結果を示した。具体的には砂浜を裸地と分類できた(図3.6a, 3.6b, 3.7a, 3.7b)が、コンクリート塗の駐車場を裸地と誤分類してしまう箇所も見られた(図3.8a, 3.8b)。

(3)は教師点の箱ひげ図から、都市と砂浜の分類に有用な特徴量であることが示された(図3.5)。また(3)は砂浜を概ね分類できたが(図3.6c),一部で砂浜を都市と誤分類してしまった(図3.7c)。一方でコンクリート塗の駐車場は都市と正しく分類できた(図3.8c)。

3.5 考察

(1)および(2)でコンクリート塗の駐車場を裸地と誤分類してしまったのは、駐車場の各強度・各成分が裸地と似ていたためと考えられる。そのためこの2つのSAR表現は砂浜と平たい土地を区別出来ず、砂浜分類には適さないと考えた。

一方(3)で駐車場を裸地と分類できたのは、コンクリートと砂浜の偏波依存性が異なっていたためと考えられる。しかしながらアルファ角では、砂浜を都市と分類してしまう箇所も

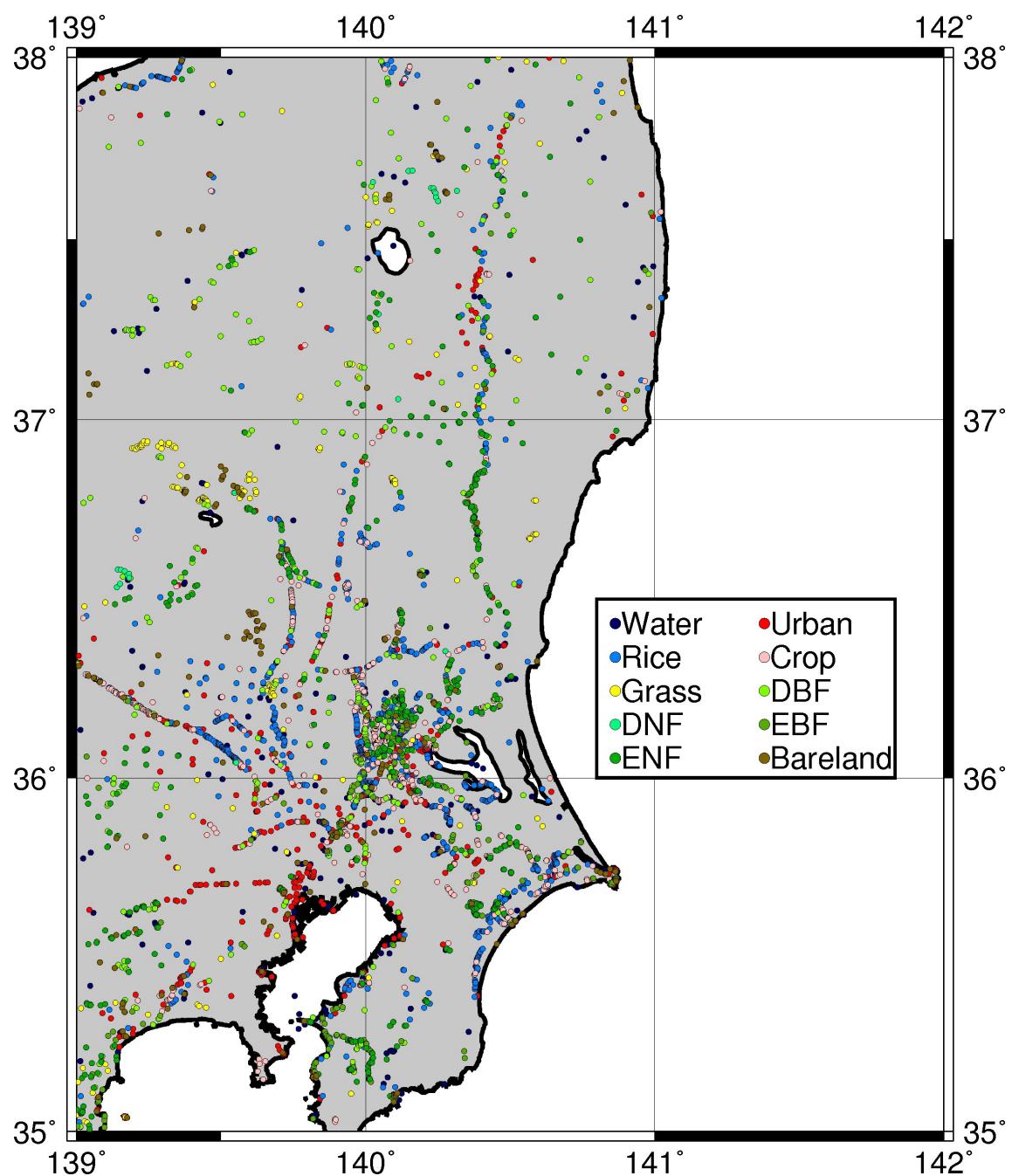


図 3.2: AVNIR2HLP の分類に使用した教師点の空間分布

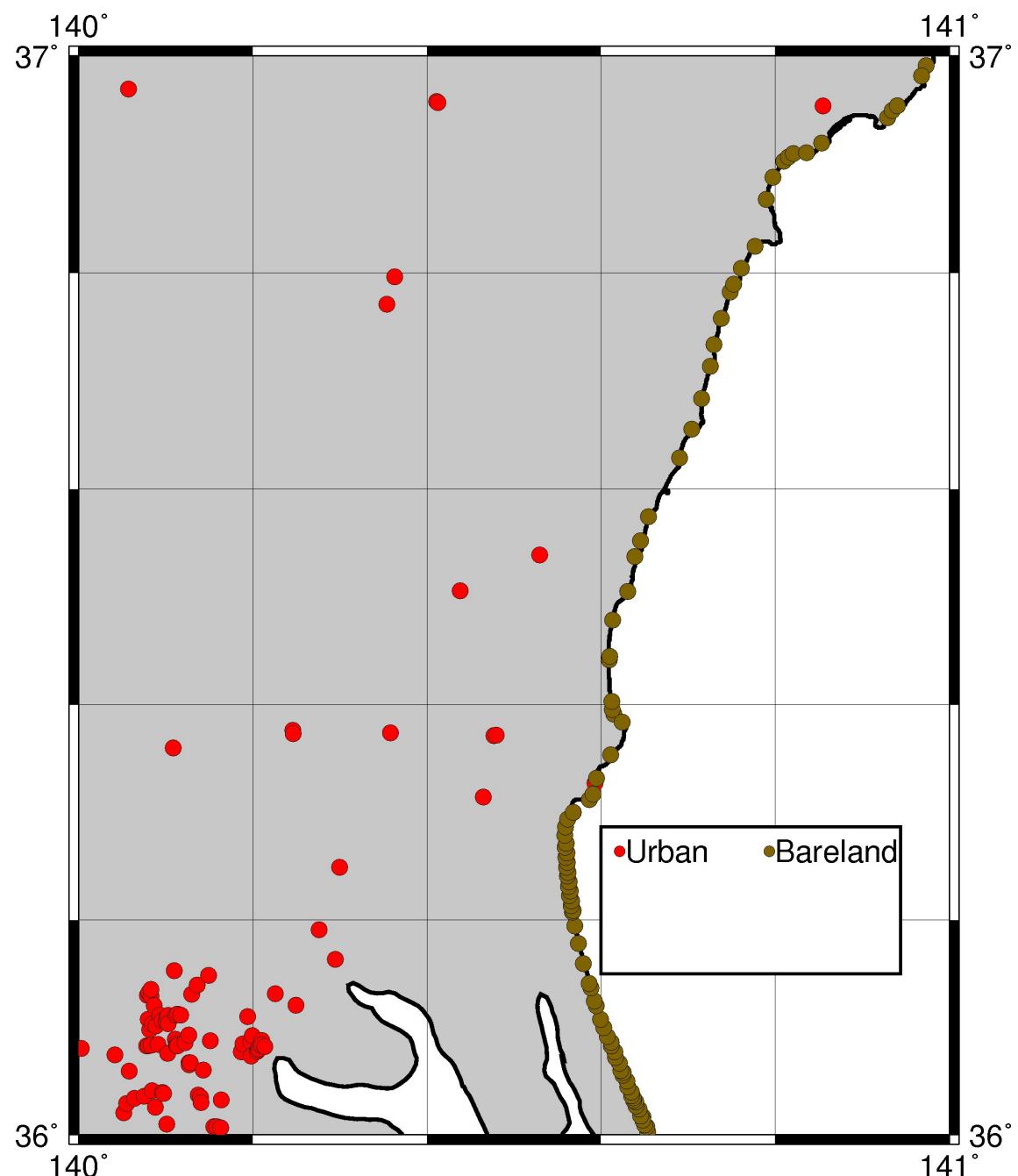


図 3.3: 多偏波 SAR の分類に使用した教師点の空間分布

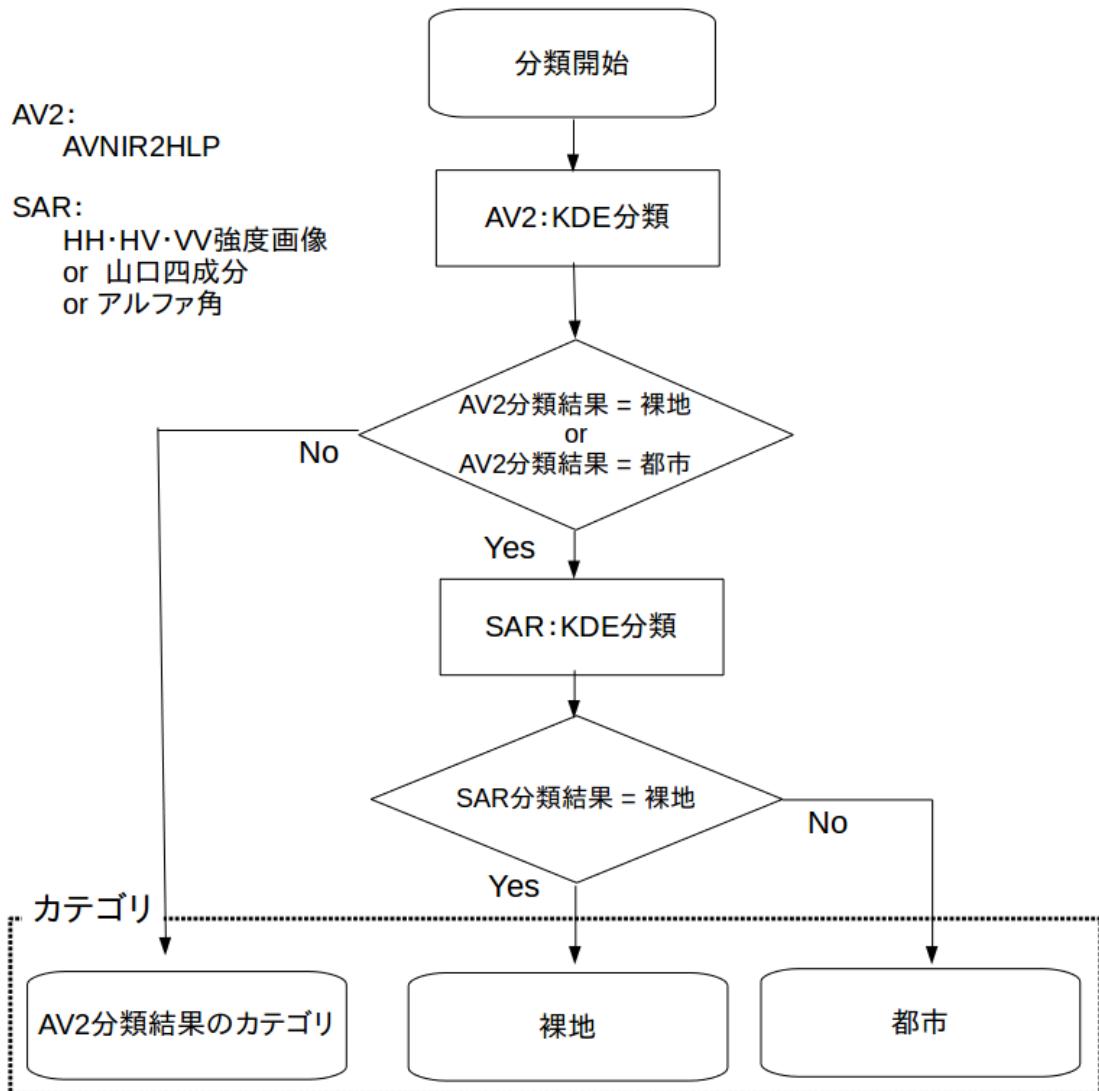


図 3.4: AVNIR2HLP と多偏波 SAR の分類のフローチャート

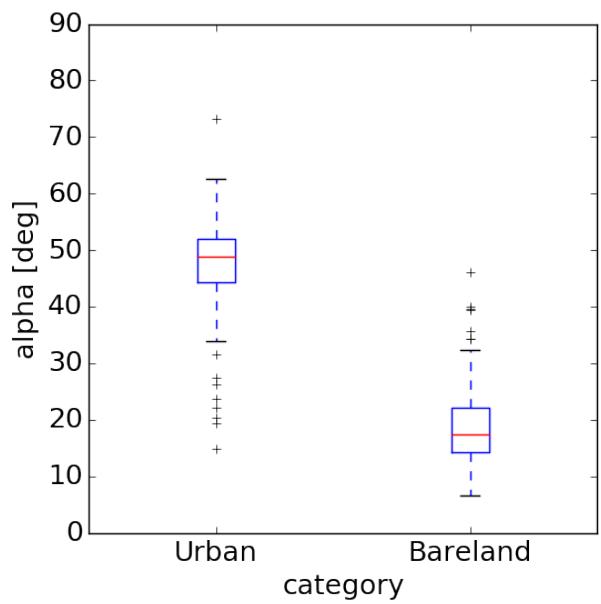
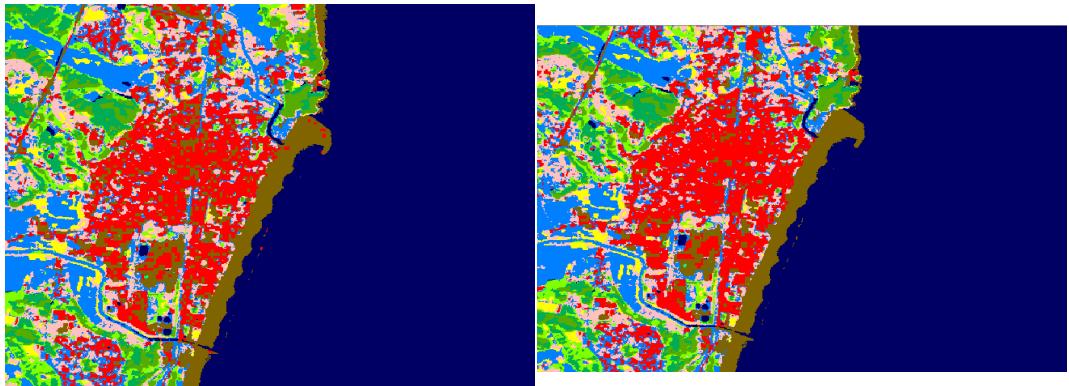


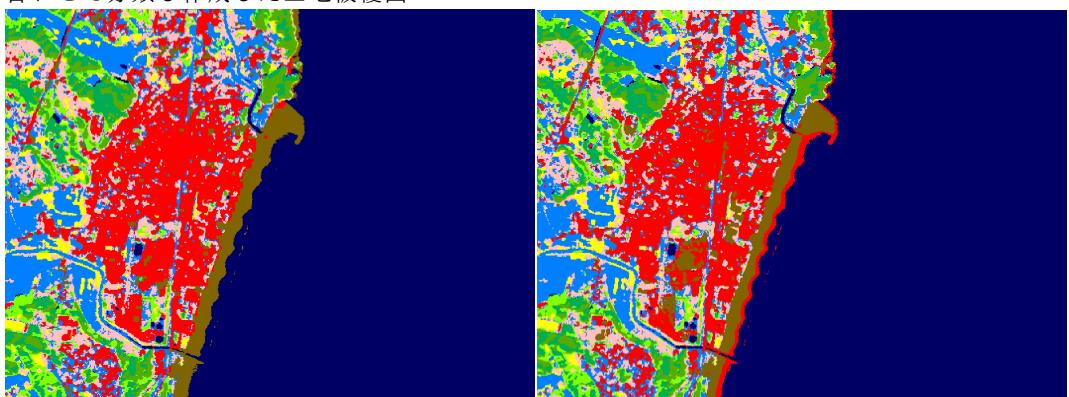
図 3.5: 都市・砂浜の教師点におけるアルファ角の箱ひげ図

見られた（図 3.7）．これは偏波強度が小さい箇所であった．そのような箇所でアルファ角を計算してしまうと，ノイズに影響されてアルファ角が突発的に大きくなってしまう．その結果，都市と誤分類したと考えられる．今後は偏波強度がどれくらい小さい時に，アルファ角の値が高くなってしまうかを検討する必要がある．



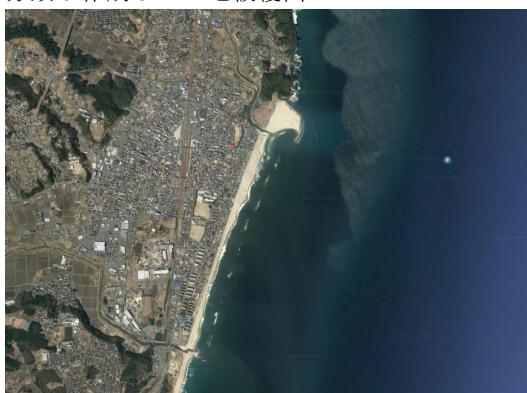
(a) HH・HV・VV強度とAVNIR2HLPを組み合わせて分類し作成した土地被覆図

(b) 山口四成分とAVNIR2HLPを組み合わせて分類し作成した土地被覆図



(c) アルファ角とAVNIR2HLPを組み合わせて分類し作成した土地被覆図

(d) AVNIR2HLPを分類し作成した土地被覆図



(e) Google Satellite画像（地図データ:Google, DigitalGlobe）

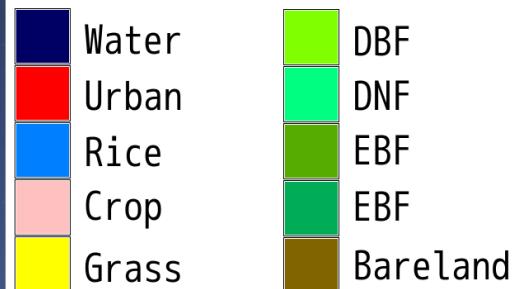
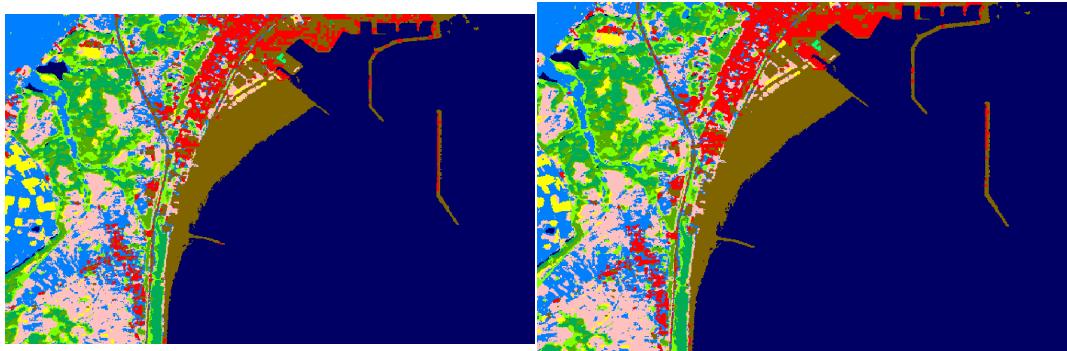
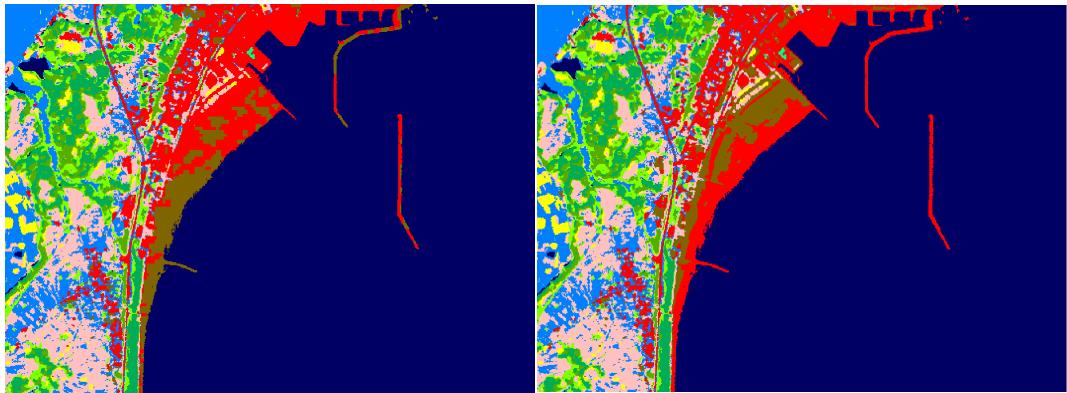


図 3.6: 高戸前浜海岸の土地被覆図と衛星画像



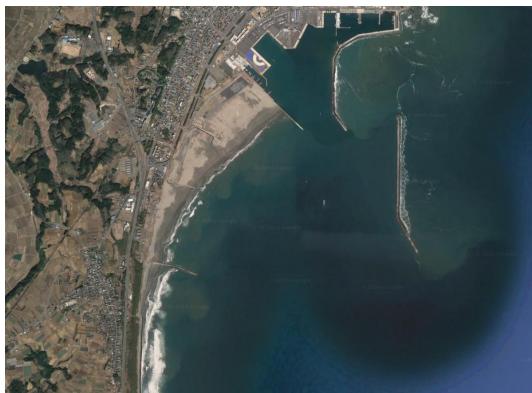
(a) HH・HV・VV 強度と AVNIR2HLP を組み合わせて分類し作成した土地被覆図

(b) 山口四成分と AVNIR2HLP を組み合わせて分類し作成した土地被覆図



(c) アルファ角と AVNIR2HLP を組み合わせて分類し作成した土地被覆図

(d) AVNIR2HLP を分類し作成した土地被覆図



(e) Google Satellite 画像（地図データ:Google, DigitalGlobe）

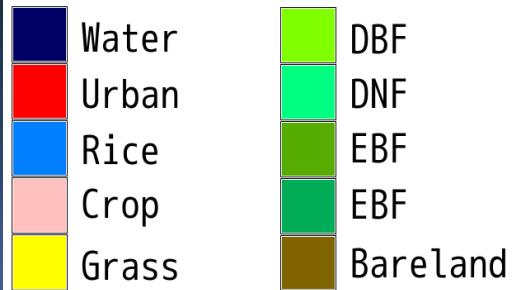


図 3.7: 大洗サンビーチの土地被覆図と衛星画像

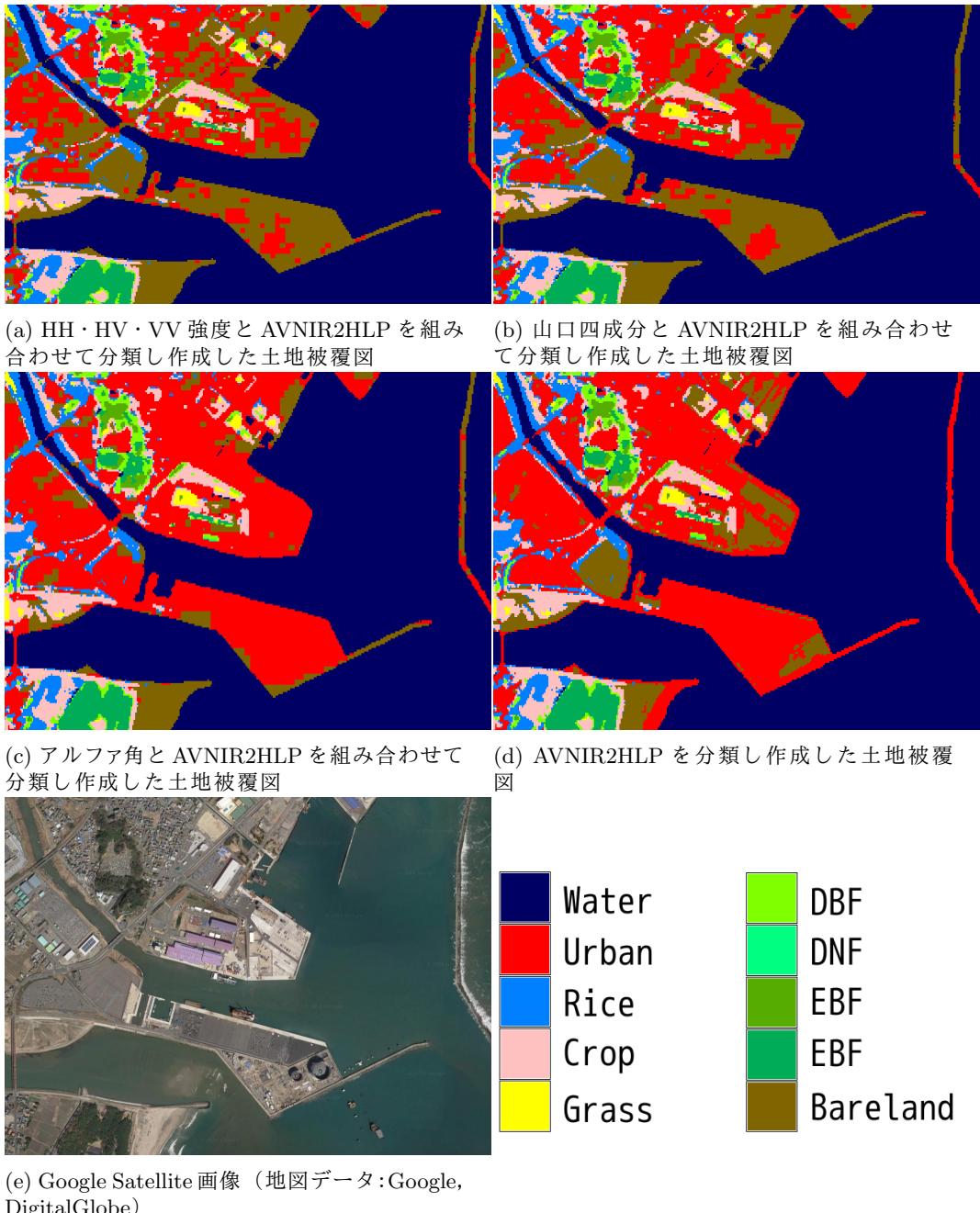


図 3.8: 日立港周辺の土地被覆図と衛星画像

第4章 多時期の二偏波SARを用いた森林・非森林マップの作成

4.1 背景

JAXA ではこれまで ALOS/PALSAR (または ALOS2/PALSAR2) の二偏波 (HH 偏波, HV 偏波) の閾値に基づいて森林・非森林マップ (Forest or Non-Forest Map; 以下 FNF マップ) を作成してきた [10]. この分類の問題点は地域ごとに閾値を決定する必要がある点である.

機械学習の教師あり学習の手法を用いると、地域ごとに閾値を設定することなく分類できる. 第1章で述べたように、JAXA は KDE 分類という教師あり学習の手法で日本域土地被覆図を作成してきた. FNF マップの作成にそこで使用していた KDE 分類・教師点を用いれば、教師あり学習と閾値に基づく分類を比較できる.

4.2 目的

多時期の二偏波 SAR と KDE 分類を組み合わせて、FNF マップを作成することを目的とした. また作成した FNF マップと JAXA の FNF マップを比較することにより、KDE 分類と閾値分類の利点・欠点を考察することも目的とした.

4.3 手法

日本全域で分類を試すには時間がかかるため、まずは小さな領域で検討する必要がある. そこで北緯 36 度東経 140 度から北緯 37 度東経 141 度の 1 度 × 1 度で、FNF マップの作成を行った. この地域を選定した理由は第3章でも述べたが、土地被覆図作成の重要なテストサイトとして情報が集約されているからである. また本研究では 2008 年から 2009 年の間は森林面積に大きな変化が無いという仮定のもと、2 年間の代表的な様子を表す FNF マップを作成した.

SAR のデータは ALOS/PLASAR が観測した二偏波 SAR 画像を用いた (表 4.1). まずははじめに SigmaSAR という JAXA の SAR 画像処理ツールを用いて、SAR 画像の再生処理を行つた (この処理は RESTEC が行った). 具体的にはオルソ処理という地表面に投影する処理を施し、 γ_0 という地形の起伏の影響を考慮した値へ変換し、DN 値 (Digital Number) として画像に格納した. そして北緯 36 度東経 140 度から北緯 37 度東経 141 度の 1 度 × 1 度の範囲で、SAR 画像を切り取った.

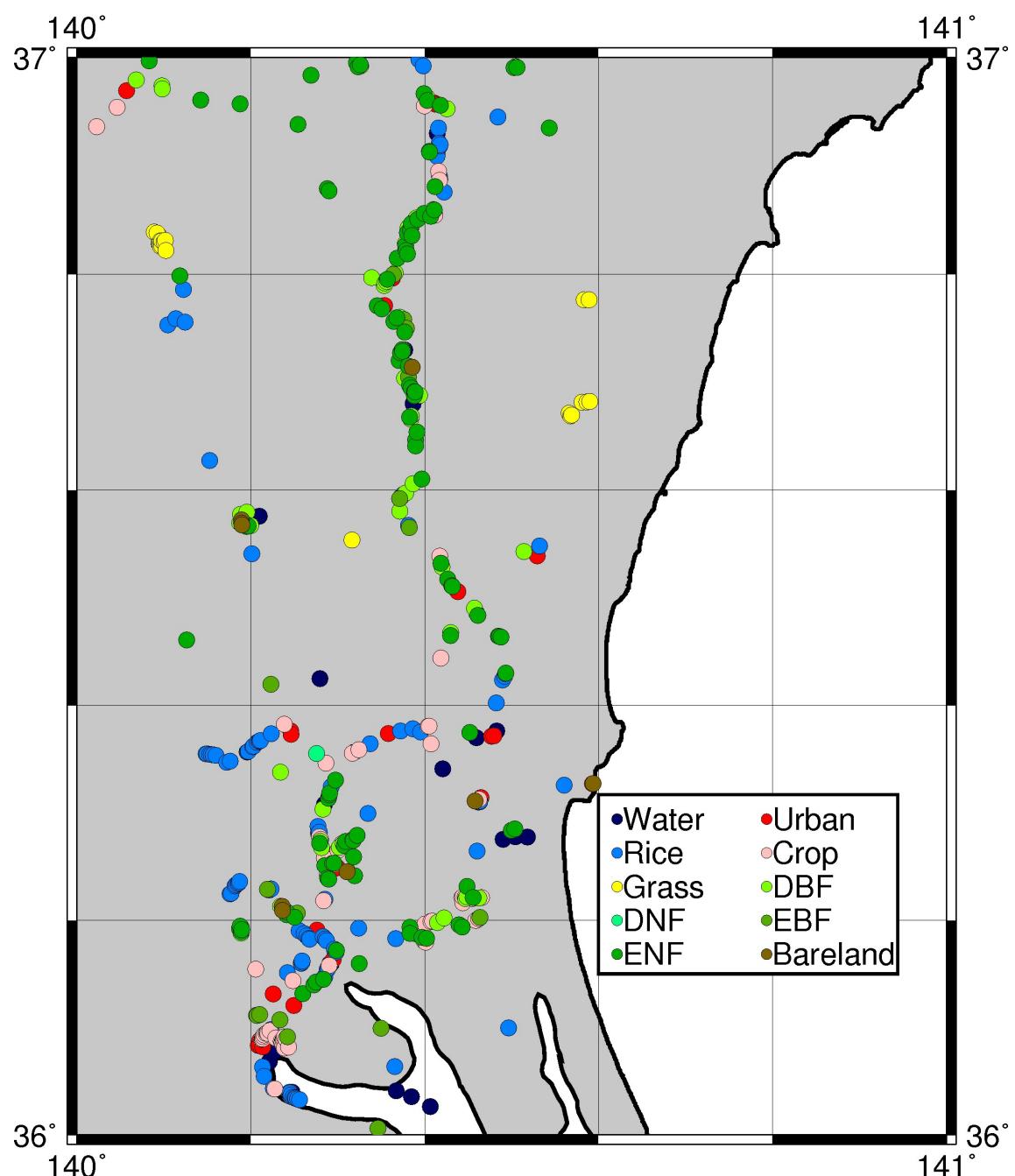


図 4.1: 二偏波 SAR の分類に用いた教師点の空間分布

表 4.1: 分類に使用した SAR 画像の概要

シーン ID	観測日	衛星進行方向	オフナディア角 (deg)	偏波
ALPSRP120120710	2008/04/26	アセンディング	34.3	HH・HV
ALPSRP140250710	2008/09/11	アセンディング	34.3	HH・HV
ALPSRP146960710	2008/10/27	アセンディング	34.3	HH・HV
ALPSRP180510710	2009/06/14	アセンディング	34.3	HH・HV
ALPSRP187220710	2009/07/30	アセンディング	34.3	HH・HV
ALPSRP193930710	2009/09/14	アセンディング	34.3	HH・HV
ALPSRP200640710	2009/10/30	アセンディング	34.3	HH・HV

さて、SAR 画像を解析する時に、スペックルノイズというノイズが問題になる。スペックルノイズは複素信号がランダムに干渉した結果生じる、空間上にランダムなノイズである [11]。これは SAR によって観測された情報を画像再生するときに施すマルチルック処理（複数の方向から見た画素を平均化する処理）のルック数を上げたり、移動窓による平均化（以下、空間平均）を施すことで低減できる。しかしノイズを低減できる代わりに、空間解像度が低くなってしまう。

近年、時間方向に平均を取ることにより空間解像度を落とすことなくスペックルノイズを低減できるという報告がなされた（本岡、私信）。ここで、時間方向に平均を取るというのは、多時期の入力画像の値を合計し、合計値を入力画像の枚数で割るという意味である。

そこで本研究ではマルチルック処理を施した画像に対して、空間平均を取る方法（式 4.1）と時間方向に平均を取る方法（式 4.2）の 2 つを試し、比較した。

$$\gamma^0 = 10 * \log_{10} < \text{DN}^2 > + \text{CF} \quad (4.1)$$

$$\gamma^0 = 10 * \log_{10} [\text{DN}^2] + \text{CF} \quad (4.2)$$

ここで DN は SAR 画像に格納されている DN 値である。また $< \text{DN}^2 >$ は（窓サイズ）×（窓サイズ）の移動窓の中で DN^2 の平均値を取ることを表し、窓サイズは比較のため 3 画素 × 3 画素、5 画素 × 5 画素の 2 つを試した。そして $[\text{DN}^2]$ は時間方向に DN^2 を平均することを表す。CF は校正係数で、-83.0 を用いた [12]。

分類は SAR 画像を 10 カテゴリで KDE 分類してから 3 カテゴリに統合する方法と、3 カテゴリで KDE 分類する方法を試した。10 カテゴリは水域（Water）・都市（Urban）・水田（Paddy）・畑地（Crop）・草地（Grass）・落葉広葉樹（DBF）・落葉針葉樹（DNF）・常緑広葉樹（EBF）・常緑針葉樹（ENF）・裸地（Bareland）とした。3 カテゴリは森林（Forest）・非森林（Non Forest）・水域（Water）とした。

教師点は第 2 章で精査したもののうち、北緯 36 度東経 140 度から北緯 37 度東経 141 度内の 1 度 × 1 度に含まれ、SAR 画像の撮影範囲に含まれるものを使用した（図 4.1）。

まずははじめに教師点を元に γ^0 を参照して分類に必要な教師データを作成した。教師データは、(i) 空間平均の場合：教師点におけるカテゴリ・HH の γ_0 ・HV の γ_0 、SAR 画像の観測日 (ii) 時間平均の場合：教師点におけるカテゴリ・HH の γ_0 ・HV の γ_0 をテキストファイルに格納した。

次に教師データを元に、KDE 分類により SAR 画像の各シーンに対して各カテゴリの尤

度を計算し、シーン毎に掛け合わせて一枚の尤度マップを作成した。ここで KDE 分類に用いた特徴空間は、(i) 空間平均の場合：HH の $\gamma_0 \cdot HV$ の $\gamma_0 \cdot t_1 \cdot t_2$ の 4 次元 (ii) 時間平均の場合：HH の $\gamma_0 \cdot HV$ の γ_0 の 2 次元である。ただし t_1, t_2 は SAR の観測日の DOY を 2 次元の単位円上で表したものである（第 2 章参照）。

そして 10 カテゴリの場合は尤度マップで尤度が最も高いカテゴリを分類結果（土地被覆図）とし、DBF・DNF・EBF・ENF を森林(Forest) に、Water を水域(Water) に、それ以外のカテゴリを非森林(Non Forest) として FNF マップを作成した。

一方 3 カテゴリの場合は尤度マップで尤度が最も高いカテゴリを FNF マップの結果とした。

4.4 結果

まずははじめに教師データの時系列箱ひげ図を図 4.2, 4.3, 4.4, 4.5 に示した。図より、水田・畑地・草地の HH の $\gamma_0 \cdot HV$ の γ_0 は季節性を持っており、水域・都市・森林・裸地は季節性がないことが分かった。また図 4.2 と 4.3, 図 4.4 と 4.5 を比較すると、窓サイズの 3 画素 × 3 画素と 5 画素 × 5 画素ではカテゴリ内の値のばらつきに大きな差が無いことが分かった。さらに図 4.2 と 4.4, 図 4.3 と 4.5 を比較すると、HH と HV はともに類似した季節性を示した。

続いて 10 カテゴリの土地被覆図から作成した FNF マップを図 4.7 に、3 カテゴリで分類した FNF マップを図 4.8 に示した。ここで分類結果の比較に、2009 年の様子を表す JAXA の FNF マップを用いた。図 4.7 と図 4.8 を比べると、3 カテゴリで分類した FNF マップの方が、10 カテゴリで作成し 3 カテゴリに統合した FNF マップより山間部の非森林のノイズが少ないことが分かった。また JAXA の FNF マップと比較すると、本研究で作成したすべてのマップで森林が過剰であった（図 4.7, 4.8）。

続いて 3 カテゴリで分類した FNF マップと JAXA の FNF マップを詳細に比較するために、霞ヶ浦と東海村の二箇所に着目し、確認した。霞ヶ浦では多時期で分類したものはハス田を非森林と正しく分類できた（図 4.9a, 4.9b）。一方で時間方向に平均を取ったものと JAXA の FNF マップは、ハス田を森林と誤分類した（図 4.9c, 4.9d）。

また東海村では空間平均・時間平均どちらも都市を森林と誤分類していた（図 4.10a, 4.10b, 4.10c）。同様に JAXA の FNF マップも都市を森林と誤分類している箇所が見られた（図 4.10d）。

4.5 考察

多時期の SAR 画像を用いることにより、JAXA の FNF マップでは森林と誤分類していたハス田を正しく分類することできた。一方で時間平均を用いて作成した FNF マップでは、ハス田を森林と誤分類した（図 4.9）。これはハス田の γ_0 には季節性があり、多時期の画像でその季節性を表すことが出来たためと考察した。時間平均を取ってしまうとその季節性がなくなってしまい、誤分類に繋がったと考えられる（図 4.6）。

また本研究で作成した FNF マップは、JAXA の FNF マップと比べて森林を過剰に検出した。そして該当箇所は都市が大半であった。これは森林と都市の γ_0 が似ており、またどちらも季節性が見られなかつたためと考察した。

さらに JAXA の FNF マップでは森林・非森林を分ける前に、オブジェクトというピクセル同士を結合した領域を作成し、分類していた。オブジェクト毎に分類することのメリットは、オブジェクトの形状の統計量や濃淡値を分類の特徴として使える点である。実際に JAXA の FNF マップでは、オブジェクト同士の境界で、HH 偏波が-1db 以上で HV 偏波が-6.5db 以上の箇所を都市として除外していた [10]。そのため本研究で作成した FNF マップより、都市を森林とする誤分類が少なかったと考えられる。

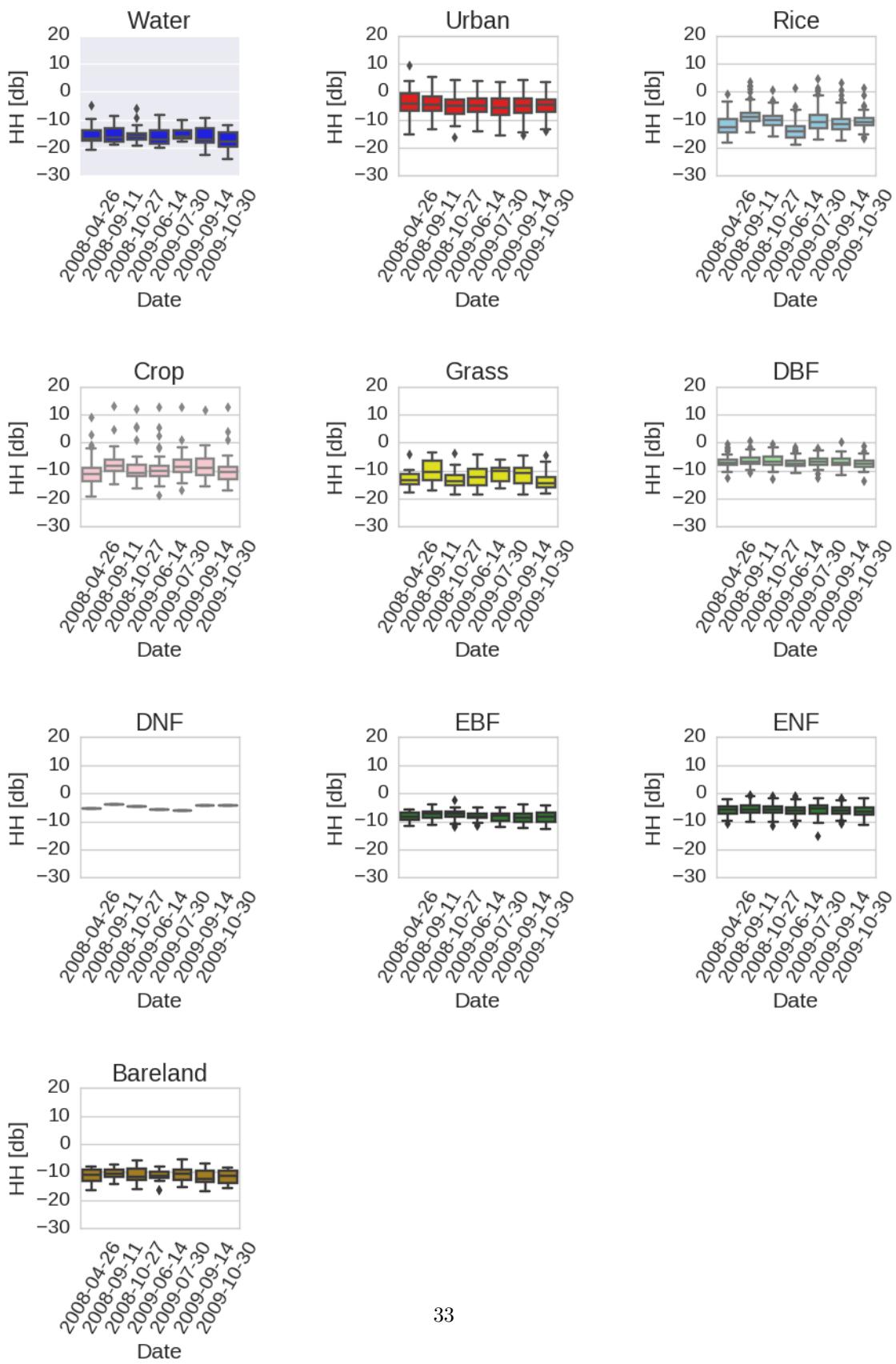


図 4.2: 教師点における HH 偏波の γ_0 の箱ひげ図. γ_0 は窓サイズ 3 画素 \times 3 画素で空間平均を取って計算した.

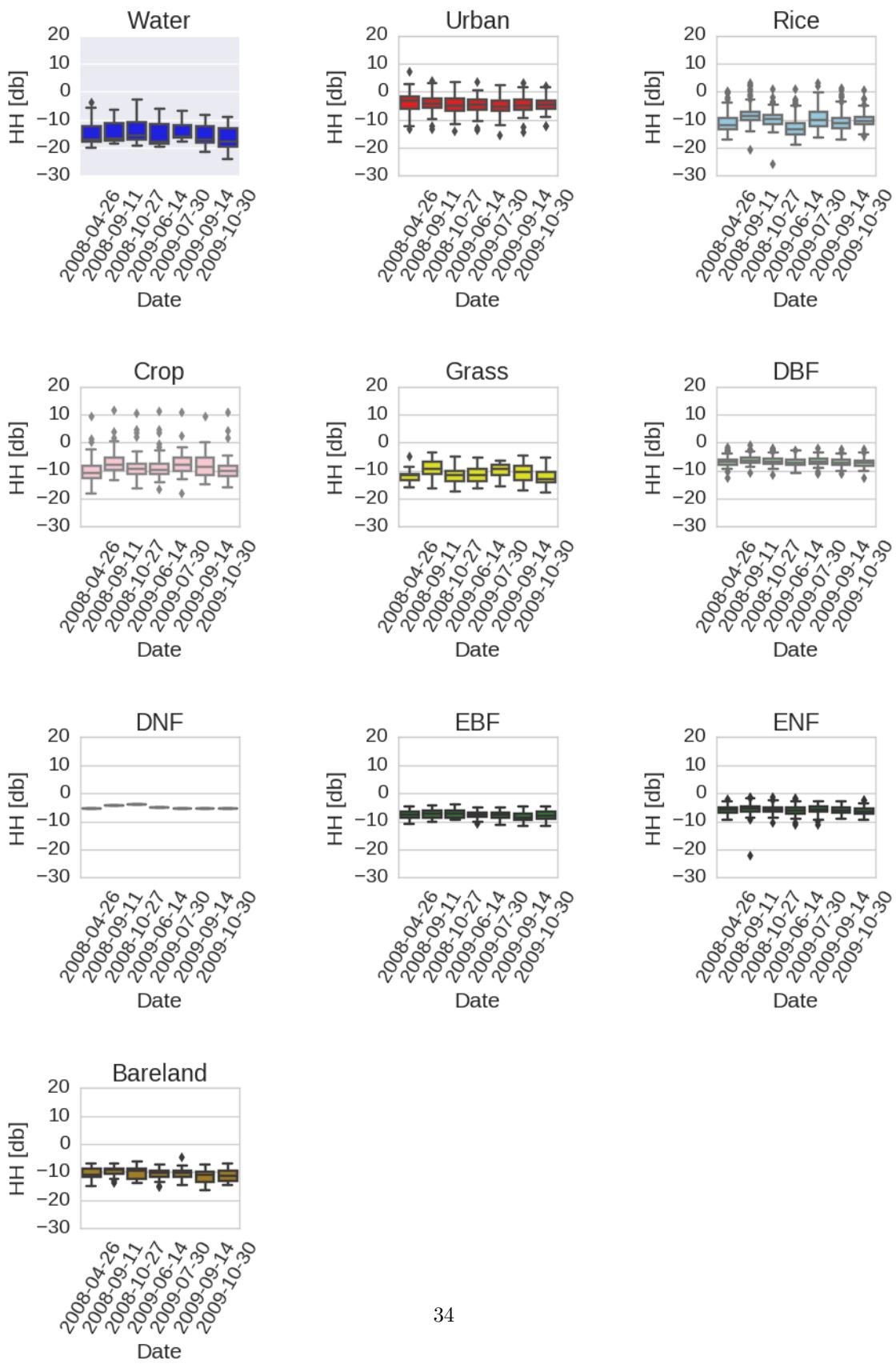


図 4.3: 教師点における HH 偏波の γ_0 の箱ひげ図. γ_0 は窓サイズ 5 画素 \times 5 画素で空間平均を取って計算した.

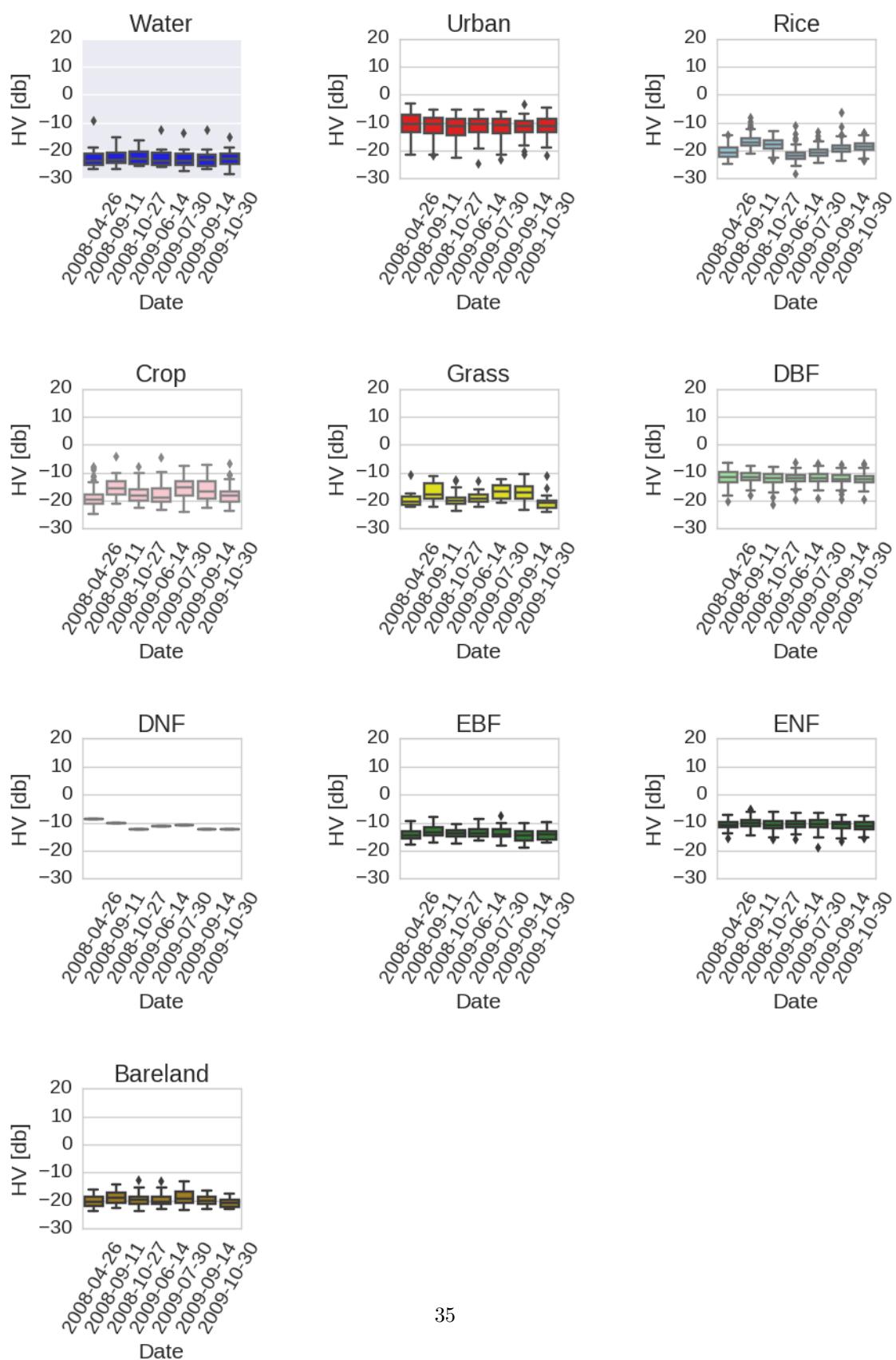


図 4.4: 教師点における HV 偏波の γ_0 の箱ひげ図. γ_0 は窓サイズ 3 画素 \times 3 画素で空間平均を取って計算した.

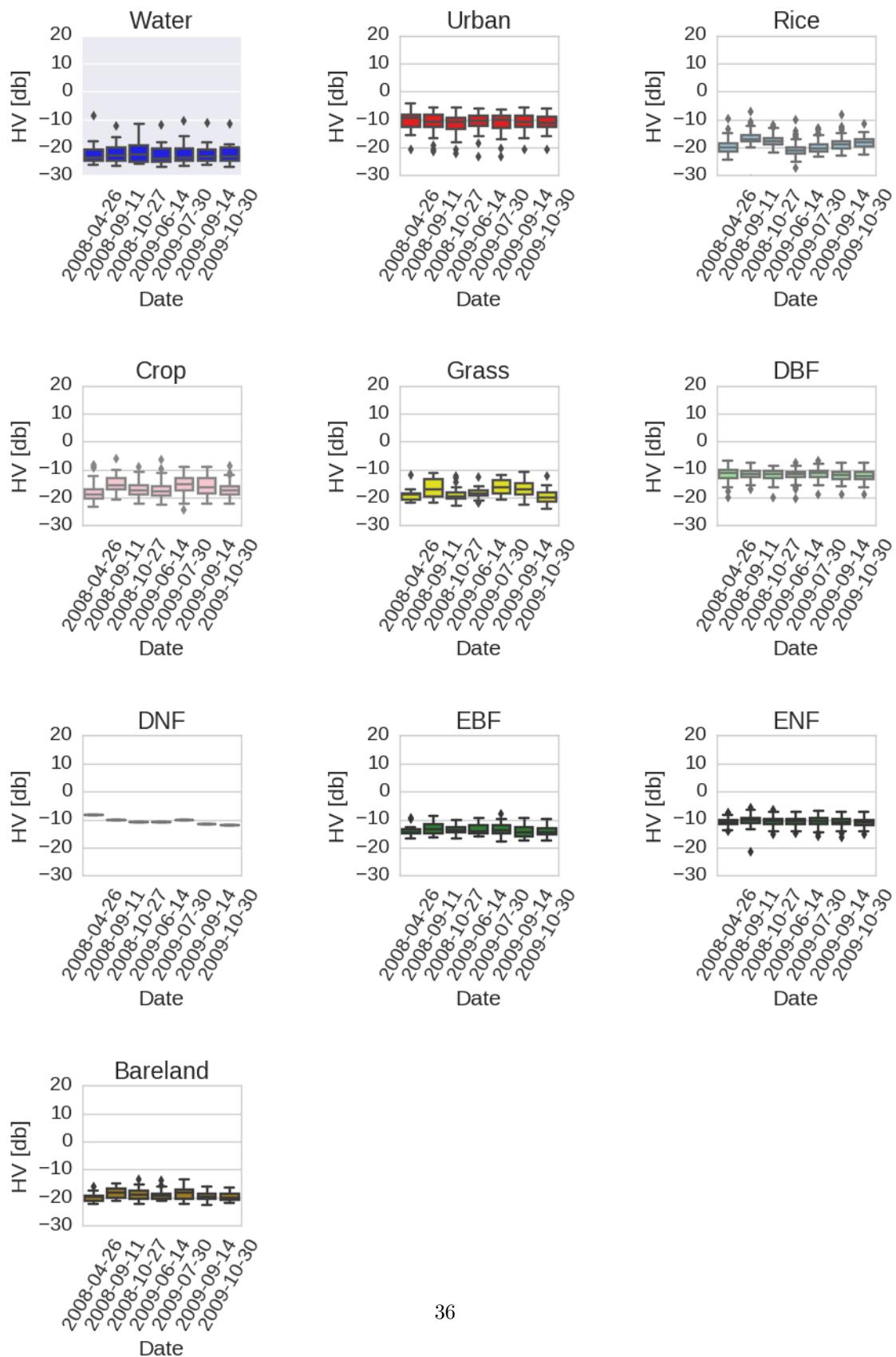
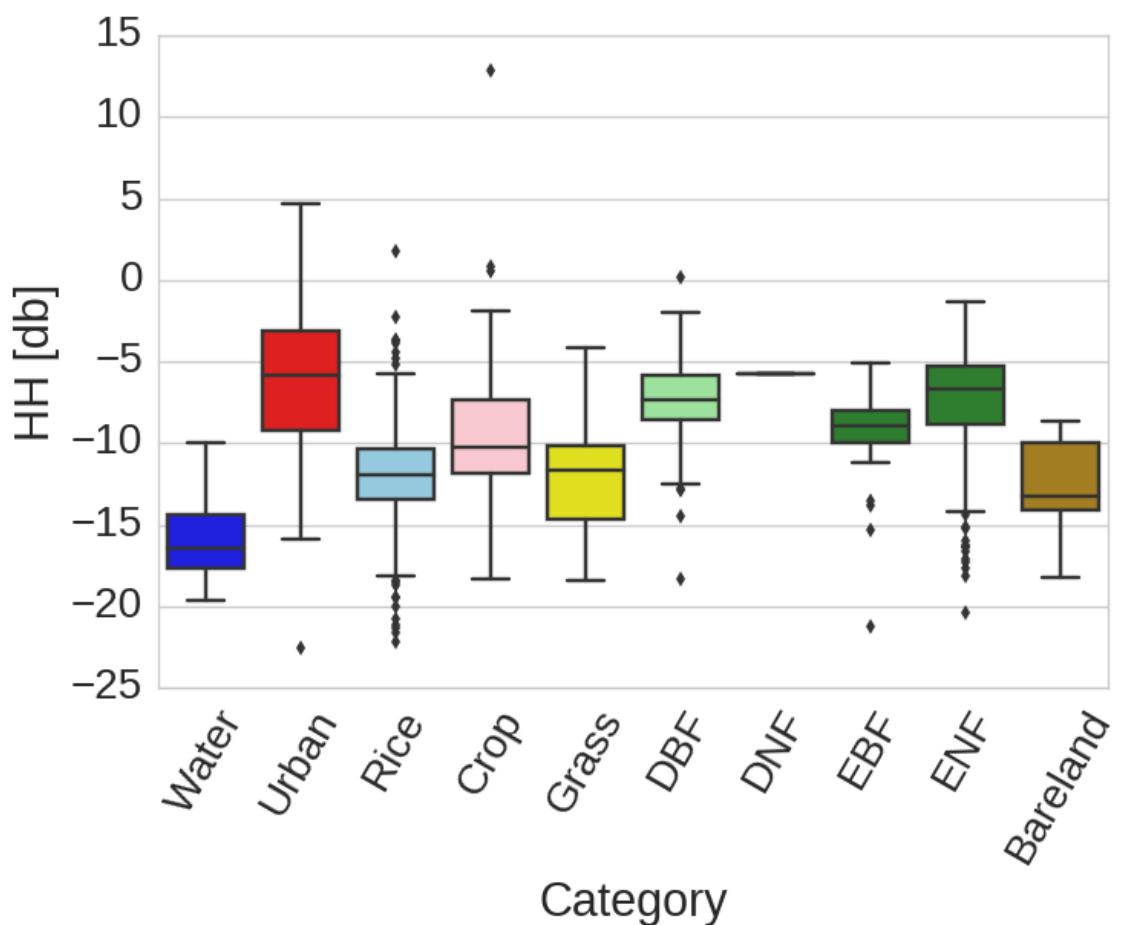
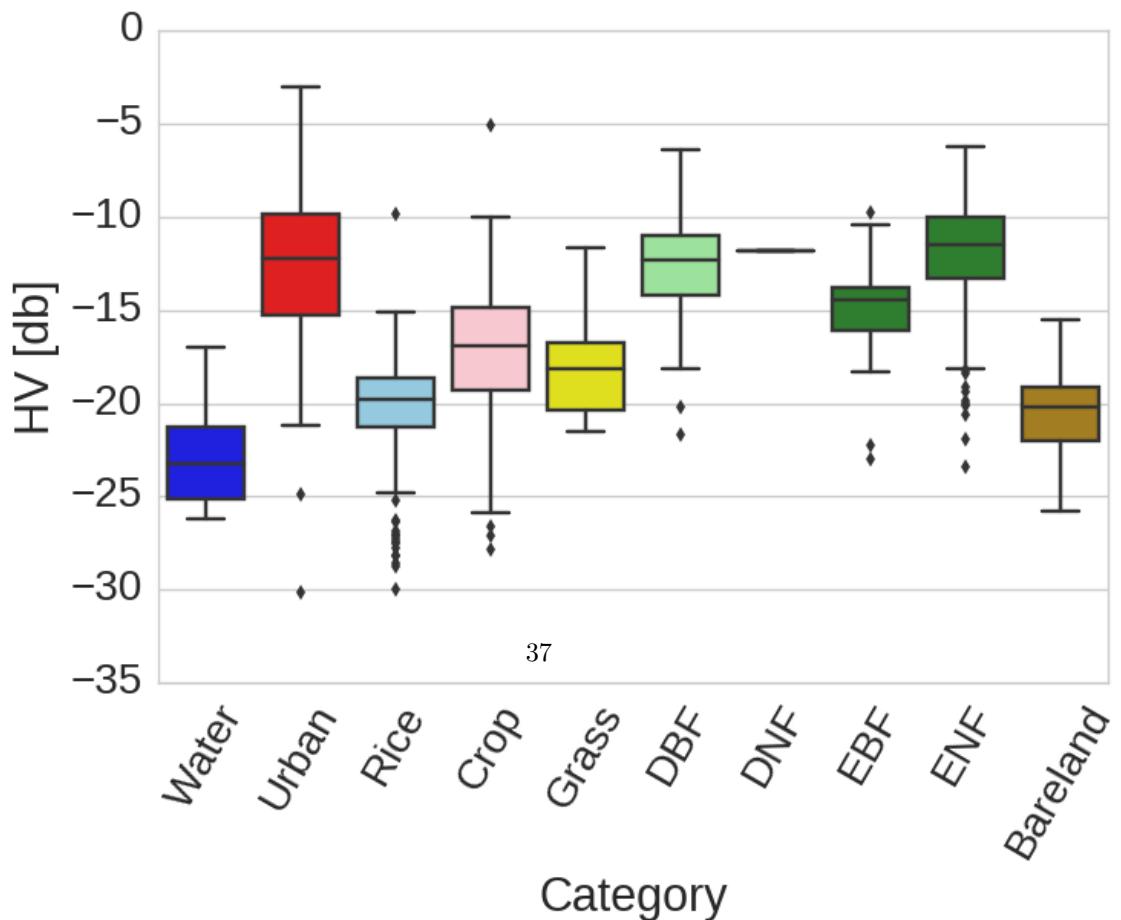


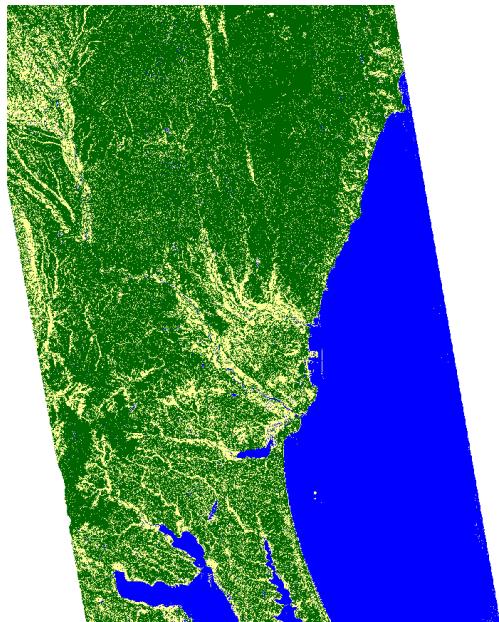
図 4.5: 教師点における HV 偏波の γ_0 の箱ひげ図. γ_0 は窓サイズ 5 画素 \times 5 画素で空間平均を取って計算した.



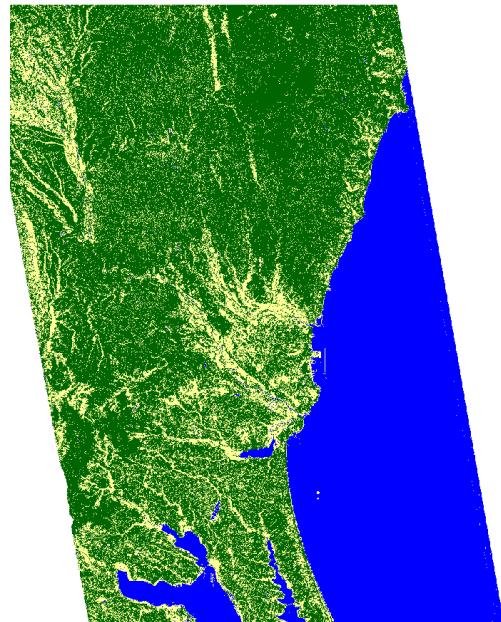
(a) HH 偏波の γ_0



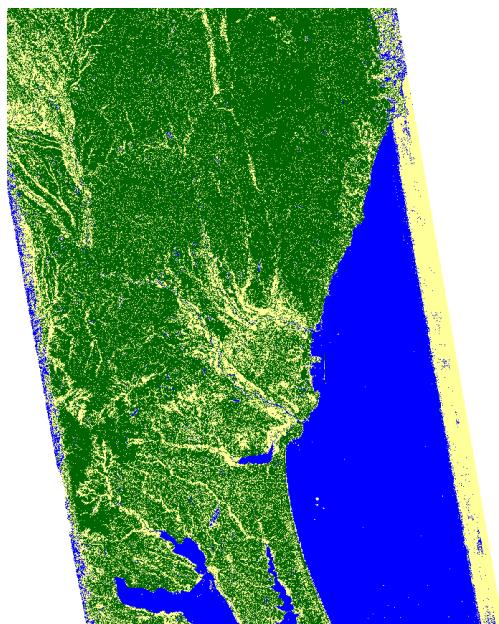
(b) HV 偏波の γ_0



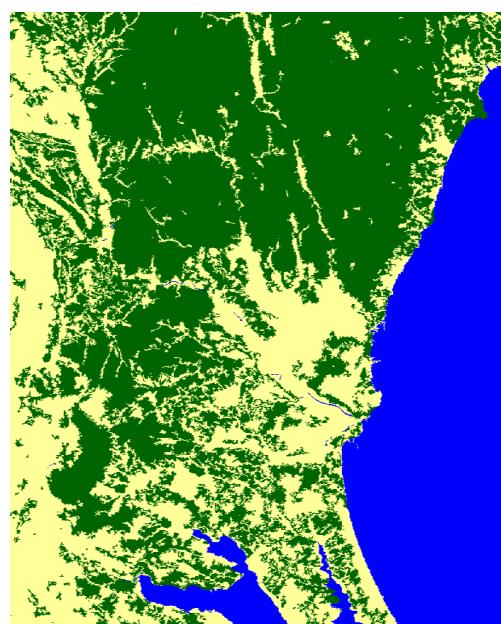
(a) 窓サイズ 3 画素 × 3 画素で空間平均を取って作成した FNF マップ



(b) 窓サイズ 5 画素 × 5 画素で空間平均を取って作成した FNF マップ



(c) 時間方向に平均を取って作成した FNF マップ



(d) JAXA の FNF マップ

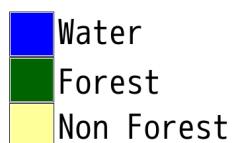
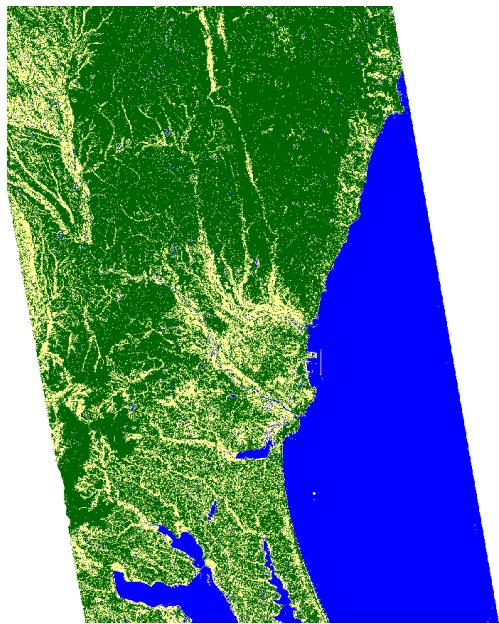
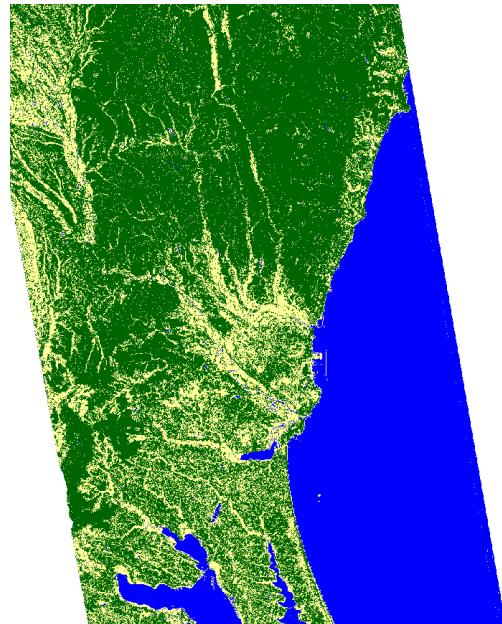


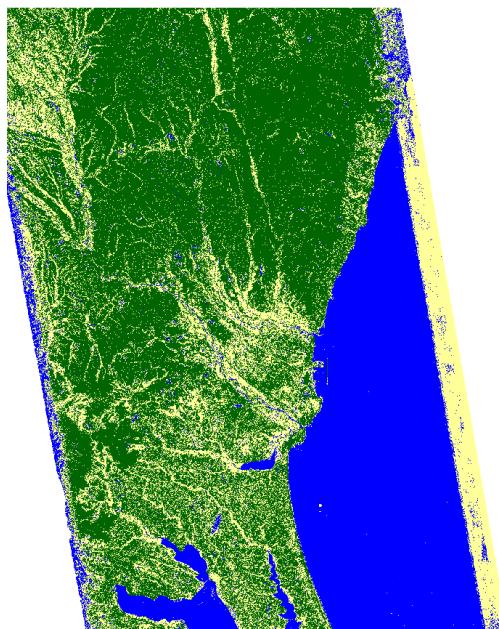
図 4.7: 多時期二偏波 SAR を用いて分類した FNF（森林・非森林）マップと衛星画像。FNF マップは 10 カテゴリから作成。



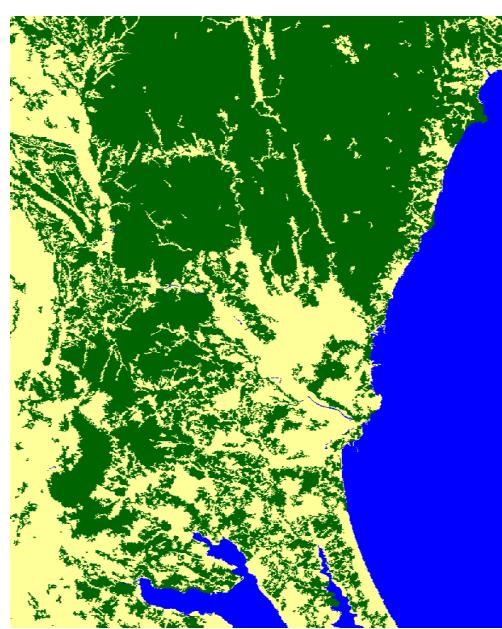
(a) 窓サイズ 3 画素 × 3 画素で空間平均を取って作成した FNF マップ



(b) 窓サイズ 5 画素 × 5 画素で空間平均を取って作成した FNF マップ



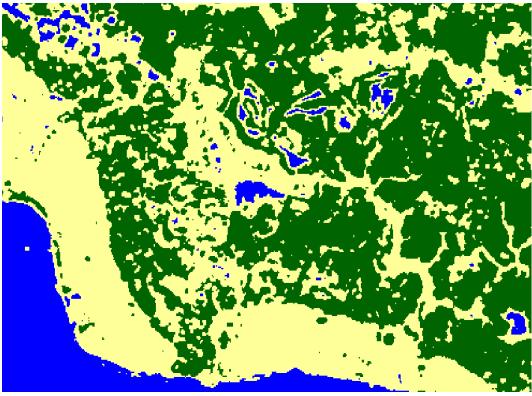
(c) 時間方向に平均を取って作成した FNF マップ



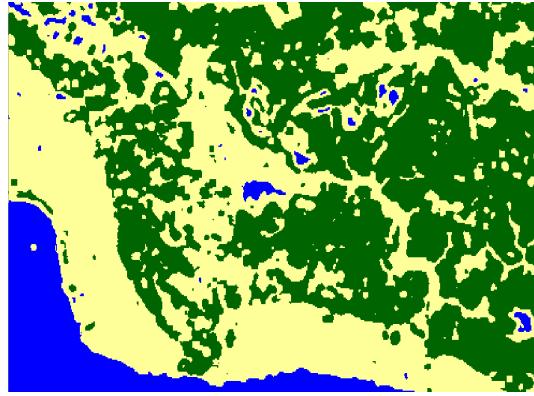
(d) JAXA の FNF マップ



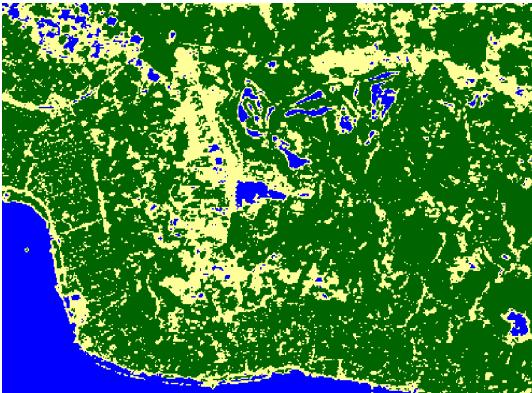
図 4.8: 多時期二偏波 SAR を用いて分類した FNF (森林・非森林) マップと衛星画像. FNF マップは 3 カテゴリから作成.



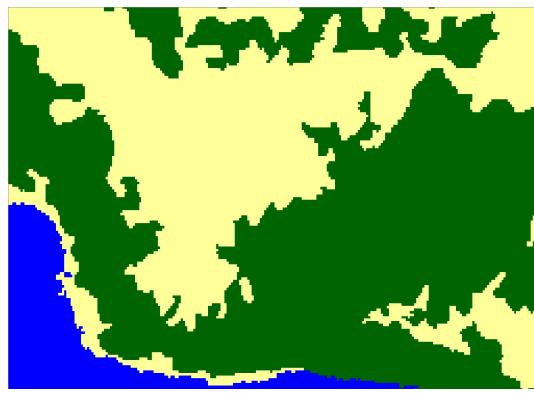
(a) 窓サイズ 3 画素 × 3 画素で空間平均を取って作成した FNF マップ



(b) 窓サイズ 5 画素 × 5 画素で空間平均を取って作成した FNF マップ



(c) 時間方向に平均を取って作成した FNF マップ



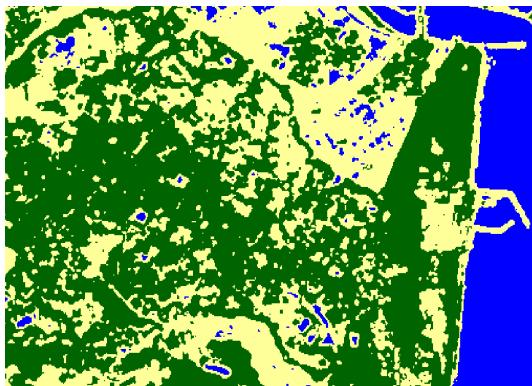
(d) JAXA の FNF マップ



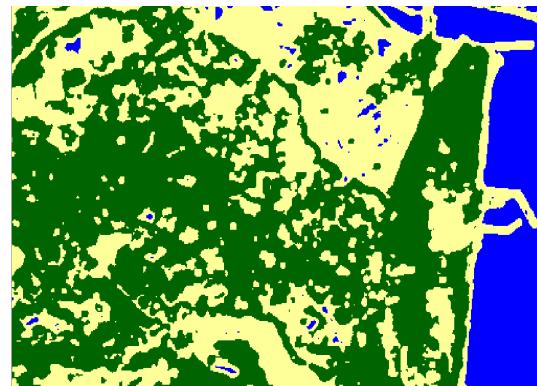
(e) Google Satellite 画像 (地図データ:Google, DigitalGlobe)

Water
Forest
Non Forest

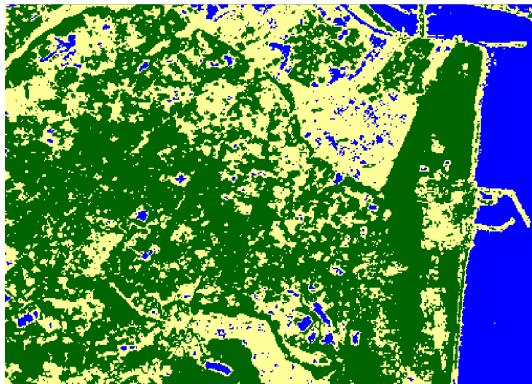
図 4.9: 多時期二偏波 SAR を用いて分類した FNF (森林・非森林) マップと衛星画像. FNF マップは 3 カテゴリから作成. 場所は霞ヶ浦周辺. 赤く囲った箇所はハス田.



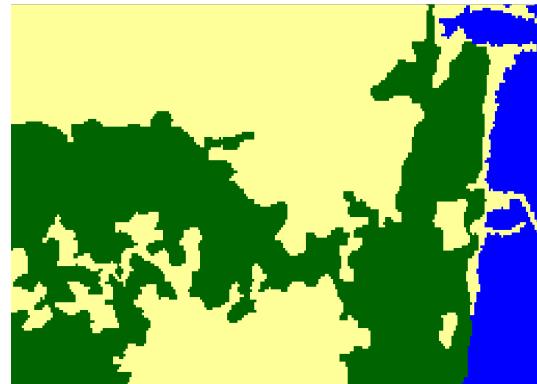
(a) 窓サイズ 3 画素 × 3 画素で空間平均を取って作成した FNF マップ



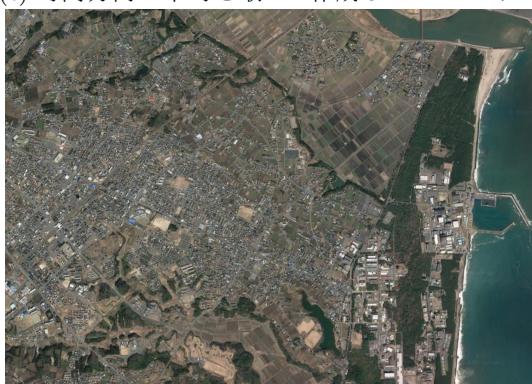
(b) 窓サイズ 5 画素 × 5 画素で空間平均を取って作成した FNF マップ



(c) 時間方向に平均を取って作成した FNF マップ



(d) JAXA の FNF マップ



(e) Google Satellite 画像 (地図データ:Google, DigitalGlobe)

Water
Forest
Non Forest

図 4.10: 多時期二偏波 SAR を用いて分類した FNF (森林・非森林) マップと衛星画像. FNF マップは 3 カテゴリから作成. 場所は東海村周辺.

第5章 結論

5.1 まとめ

入力データの精査を行うことにより、分類精度が向上することが第2章で示された。

第3章では多偏波SARと光学センサを組み合わせることにより、海岸域の砂浜の誤分類を低減できることができた。また偏波強度、山口四成分を用いた分類では駐車場を裸地と誤分類してしまうが、アルファ角は誤分類しないことが分かった。しかしアルファ角では砂地を都市と誤分類してしまう箇所も見られた。

第4章では二偏波SARを時間方向に平均を取るより、空間方向に平均を取る方が誤分類が少ないことが分かった。また多時期のSAR画像を用いることにより、ハス田などの非森林域の季節性を表すことができ、誤分類が低減することができた。さらに二偏波SARを用いたオブジェクト毎の閾値分類（JAXAのFNFマップ）より、多時期二偏波SARを用いたピクセル毎のKDE分類（本研究で作成したFNFマップ）の方が森林が過剰であることが分かった。

そして本研究の成果として、第2章で作成した土地被覆図をJAXAの正式なプロダクトとしてリリースした[13]。

5.2 今後の課題

以下に本論文から得られた課題、及び本論文で取り扱えなかった今後の課題を示す。

5.2.1 事前確率

SARを用いて土地被覆図を作成すると、森林が過剰となることが第4章で分かった。第2章でSARを先駆情報として取り入れているため、第2章で作成した土地被覆図は森林が過剰になっている可能性がある。

また第2章で作成した土地被覆図は、2006年から2011年の間の代表的な様子を表すものである。しかし第2章で用いたOpenStreetMapのデータは2016年のものである。道路の分布は頻繁に変わるために、2011年以降に出来た道路や廃止された道路が分類結果に影響している可能性がある。

続いて第2章で事前確率及びAVNIR2HLPの分類結果を統合するときに、1:1の重みで掛けあわせている。そのため事前確率の分類結果の精度が悪いと、その影響が強く出てしまう可能性がある。

5.2.2 教師点

教師点の地点の土地被覆が変化すると、土地被覆図の精度が低下することが第2章で分かった。教師点はOpen Layers の Google Satellite 画像とSACL AJを参照して作成した。しかしながらGoogle Satellite 画像では単時期の情報しか参照できず、またその画像が撮影された時期も、作成する土地被覆図の対象時期とは異なっている可能性がある。今後もOpen Layers を用いて教師点を作成する場合は、土地被覆図の対象時期をカバーするレイヤマップを作成し、それを参照しながら教師点を取得する必要がある。代替案としては、過去の衛星画像を参照できるGoogle Earth を利用することが挙げられる。

続いて本研究では教師データの取得範囲を1度×1度あるいは3度×3度のいずれかにした。一般に教師点は数が多いほど分類精度は良くなると言われている。だが広い範囲から教師点を沢山取ってしまうと、異なる衛星パスにより生じる反射率の違いや地域性の違いから誤分類の原因に繋がると考えられる。この教師点の取得範囲の最適解は未だ無い。

第2章・第3章・第4章で使用した教師点は、道路沿いや鉄道沿いに多かった。これは教師点の作成に、現地踏破写真に基づくSACL AJというサイトを用いたためである。現地踏破に基づくと、アクセスしやすい箇所に教師点が集中してしまう。教師点の空間分布を考慮すると、使用する教師点のサンプリング手法を検討する必要がある。

5.2.3 カテゴリ

第2章で作成した土地被覆図は10カテゴリと主要な項目にとどめた。これは利用者が自分の目的にそってチューニングしやすくするためである。しかし分類結果や地域性を考慮すると、河川の草地（抽水植物）や茶畠などのカテゴリの追加も考えられる。またFAOのLCCS（Land Cover Classification System）[1]への準拠も検討事項である。

第2章で作成した土地被覆図では、以下の誤分類が見られる。

1. ビニールハウスが都市になる
2. スキー場が水田になる
3. 河原の裸地が都市・水田になる
4. 埋立地の畑問題
5. 水田 or 森林問題（例：N26E127）
6. 道路が畑などミックスピクセルになっている

5.2.4 手法・アルゴリズム

KDE分類では分類結果の各シーンを統合するときに、ベイズ推定に基づく掛け合わせを行っている。シーン数が多いと特定のカテゴリの値が非常に大きくなり、他のカテゴリが0になってしまう問題がある。

5.2.5 検証

第2章では予め取得した検証点における混合行列を用いて検証を行った。予め取得した検証点を用いるメリットとしては、土地被覆図を改良した時に改良前後で精度評価を素早く行える点である。しかしカテゴリごとに検証点の数に偏りがあるため、母集団で分布数が少ないカテゴリ（茨城県における落葉針葉樹など）の評価が安定しない。

そこで層別ランダムサンプリングという手法を用いると、カテゴリの空間分布を考慮した検証が行える。デメリットはランダムにばらまいた点のカテゴリの判別が、判別する人によって異なる恐れがあることである。

謝辞

本研究で使用した AVNIR-2 データは、JAXA・筑波大学共同研究契約の枠組みで宇宙航空研究開発機構地球観測研究センター (JAXA/EORC) で処理された高次成果物である。

本研究を行う上で、多くの方々にお世話になりました。RESTEC の小林健一郎さんにはプログラムのバグ探しから研究のサポートまで幅広く協力していただきました。また JAXA の田殿武雄さん、大木真人さんには PALSAR モザイクデータの使用を快く許可していただきました。RESTEC バイトの平山智弘くん・内田圭祐くん・高橋朋也くん・前原大成くんには、分類結果のチェックや教師点の取得、教師点の精査を手伝ってもらいました。

筑波大学流域管理研究室の皆様には公私ともにお世話になりました。特に奈佐原顕郎先生には研究で行き詰った時のアドバイスから日本語の添削まで丁寧に指導していただき、心より感謝いたします。

最後に本研究は GDAL/OGR, GRASS GIS, QGIS を始めとした多くのオープンソース・ソフトウェアによって実現しました。これらのソフトウェアのコミュニティの皆様には深く感謝いたします。

参考文献

- [1] FAO (2000) : LAND COVER CLASSIFICATION SYSTEM (LCCS) : CLASSIFICATION CONCEPTS AND USER MANUAL, <http://www.fao.org/docrep/003/x0596e/x0596e00.HTM>.
- [2] 片木仁, 小林健一郎, 田殿武雄, 奈佐原顕郎 (2016): JAXA 日本域高解像度土地利用・土地被覆図の高精度化 (Version 16.09), 日本リモートセンシング学会第 61 回学術講演会論文集, pp.87-88.
- [3] 橋本秀太郎, 田殿武雄, 小野里雅彦, 堀雅裕 (2014): 多時期光学観測データを用いた高精度土地被覆分類手法の開発, 日本リモートセンシング学会誌, 34 (2), pp.102-112.
- [4] Russell G.Congalton, Jianyu Gu, Kamini Yadav, Prasad Thenkabail and Muthu Ozdogan (2014): Global Land Cover Mapping: A Review and Uncertainty Analysis. *Remote Sens.*, 6, pp.12070-12093.
- [5] © OpenStreetMap Contributors
- [6] 農林水産省: 農林水産関係市町村別統計, 平成 26 年産 水稲 北海道
- [7] 山口芳雄 (2007): レーダポーラリメトリの基礎と応用—偏波を用いたレーダリモートセンシング—, 電子情報通信学会.
- [8] S.R.Cloude and E.Pottier (1997): An entropy based classification scheme for land applications of polarimetric SAR, *IEEE Trans. Geosci. Remote Sens.*, 35, 1, pp.68-78.
- [9] Jong-Sen Lee, Thomas L. Ainsworth, John P.Kelly, and Carlos Lopez-Martinez (2008): Evaluation and Bias Removal of Multilook Effect on Entropy/Alpha/Anisotropy in Polarimetric SAR Decomposition. *IEEE Trans. Geosci. Remote Sens.*, vol.46, no.10, october 2008.
- [10] Masanobu Shimada, Takuya Itoh, Takeshi Motooka, Manabu Watanabe, Shiraishi Tomohiro, Rajesh Thapa, and Richard Lucas (2014) : New Global Forest/Non-forest Maps from ALOS PALSAR Data (2007-2010), *Remote Sensing of Environment*, 155, pp.13-31.
- [11] 日本リモートセンシング学会編 (2011) : 基礎からわかるリモートセンシング, 理工図書株式会社, p278.
- [12] JAXA (2009) : PALSAR の校正処理の変更について, http://www.eorc.jaxa.jp/about/distribution/info/20090109jp_3.html.

[13] JAXA (2016) : 日本域高解像度土地利用土地被覆図 2016年9月リリース版 (バージョン 16.09) , http://www.eorc.jaxa.jp/ALOS/lulc/jlulc_jpn.htm