



Title	合成開口レーダを利用した圃場情報の取得に関する研究
Author(s)	山谷, 祐貴
Degree Grantor	北海道大学
Degree Name	博士(農学)
Dissertation Number	甲第14374号
Issue Date	2021-03-25
DOI	<a href="https://doi.org/10.14943/doctoral.k14374">https://doi.org/10.14943/doctoral.k14374</a>
Doc URL	<a href="https://hdl.handle.net/2115/96747">https://hdl.handle.net/2115/96747</a>
Type	doctoral thesis
File Information	Yamaya_Yuki.pdf



合成開口レーダを利用した  
圃場情報の取得に関する研究

北海道大学 大学院農学院

共生基盤学専攻 博士後期課程

山 谷 祐 貴



# 目次

1	緒言	1
1.1	研究の背景	1
1.2	リモートセンシングの概要	1
1.3	合成開口レーダと偏波	3
1.4	散乱成分分解	4
1.5	SAR データ利用に関する既往研究	5
1.5.1	作付作物の分類	5
1.5.2	生育状況の推定	7
1.6	本研究の目的	9
2	Xバンド SAR データを使用した機械学習アルゴリズムによる作付作物の分類	13
2.1	背景と目的	13
2.2	研究方法	14
2.2.1	対象地域および対象作物	14
2.2.2	使用データおよび処理方法	15
2.2.3	解析方法	16
2.2.4	解析の流れ	18
2.3	結果と考察	19
2.3.1	決定木の数の検討結果	19
2.3.2	全変数利用時の分類の精度	20
2.3.3	要素数別の分類の精度	21
2.3.4	時期数別の分類の精度	23
2.3.5	作付マップの作成	25
2.4	まとめ	25

3	XバンドおよびCバンド SAR データを併用した機械学習アルゴリズムによる作物分類の高精度化・効率化に関する検討	29
3.1	背景と目的	29
3.2	研究方法	30
3.2.1	使用データおよび処理方法	30
3.2.2	解析方法	32
3.2.3	解析の流れ	35
3.3	結果と考察	36
3.3.1	現地調査結果	36
3.3.2	決定木の数の検討結果	42
3.3.3	RADARSAT-2 データを使用した分類	42
3.3.4	RADARSAT-2 データと TerraSAR-X データとを組み合わせた分類	48
3.3.5	使用する時期数の検討	52
3.4	まとめ	60
4	SAR を利用した高精度な作付作物分類のための機械学習アルゴリズムと衛星データの選定	63
4.1	背景と目的	63
4.2	研究方法	65
4.2.1	使用データおよび処理方法	65
4.2.2	解析方法	67
4.2.3	解析の流れ	69
4.3	結果と考察	70
4.3.1	機械学習アルゴリズムの比較	70
4.3.2	RADARSAT-2 データと TerraSAR-X データとの分類結果の比較	76
4.4	まとめ	86

5	CバンドおよびLバンド SAR データを併用した機械学習アルゴリズムによる作物分類の効率化に関する検討.....	89
5.1	背景と目的.....	89
5.2	研究方法.....	91
5.2.1	使用データおよび処理方法.....	91
5.2.2	解析方法.....	93
5.2.3	解析の流れ.....	93
5.3	結果と考察.....	95
5.3.1	1 時期を使用した分類でのバンド間の比較.....	95
5.3.2	多時期を使用した分類でのバンド間の比較.....	97
5.3.3	一部の RADARSAT-2 を PALSAR-2 で代替した分類.....	104
5.4	まとめ.....	106
6	Cバンド SAR および光学センサを使用した水稻穂含水率の推定.....	109
6.1	背景と目的.....	109
6.2	研究方法.....	111
6.2.1	対象地域および対象作物.....	111
	.....	113
6.2.2	使用データおよび処理方法.....	113
6.2.3	解析方法.....	114
6.2.4	解析の流れ.....	116
6.3	結果と考察.....	117
6.3.1	現地調査結果.....	117
6.3.2	SAR データを使用した水稻穂含水率の推定モデル作成エラー! ブックマークが定義されていません。	
6.3.3	SAR データおよび光学センサデータを併用した水稻穂含水率の推定モデル作成.....	125
6.3.4	SAR データを使用した水稻穂含水率の推定モデルの検証.....	128
6.4	まとめ.....	132

7	Cバンド SAR およびダミー変数を使用した水稻穂含水率の推定	135
7.1	背景と目的	135
7.2	研究方法	136
7.2.1	使用データおよび処理方法	136
7.2.2	解析方法	137
7.2.3	解析の流れ	138
7.3	結果と考察	139
7.3.1	SAR 変数とダミー変数を使用した水稻穂含水率の推定モデル作成	139
7.3.2	SAR 変数とダミー変数を使用した水稻穂含水率の推定モデルの検証	153
7.3.3	SAR 変数とダミー変数を使用した水稻穂含水率の推定マップの作成	154
7.4	まとめ	156
8	結言	159
	謝辞	163
	引用文献	165

# 1 緒言

## 1.1 研究の背景

農業地域における圃場ごとの作物に関する情報は、作目ごとの作付面積の調査、生産量の予測、生育状態の把握、災害発生時の被害調査など、様々な用途に供せられており、今後も継続的かつ長期的、広域的に把握していくことが重要である。例えば農林水産省では、作物統計調査として職員や統計調査員による耕地面積および作付面積の実測調査が行われている（農林水産省、2020）。しかし、現在人の手によって行われている現地調査は、人手不足により調査が縮小される懸念もあり、作付状況に関する情報の十分な把握が、今後は困難となるおそれがある。したがって、人的コストの削減を進める一方で、圃場ごとの作物に関する情報を詳細に把握していく手法の確立が必要であり、これらの両立が日本の農業地域における課題の一つとなっている。

また、地球観測戦略における課題として、食料の収穫状況や収量を効果的かつ継続的に監視することが挙げられており、これら農業生産に関する情報は先進国、発展途上国ともに今後重要視されていくと考えられる（辻野、2014）。すなわち、農業に関する情報を広範囲かつ効率的に取得することは、農業分野全体の産業競争力強化をもたらすための課題となると考えられる（高度情報通信ネットワーク社会推進戦略本部、2014）。

これらの課題の解決策として、従来の現地調査に代わる手法の確立が不可欠である。その中でも衛星リモートセンシング技術は、農業分野において対象物の状態を非破壊で計測する手法として期待されており（齋藤ら、2017）、作付品目の特定や作付面積の把握、作物生育の把握などにますます利用されることが期待されている（野口、2016）。したがって、衛星観測データの効率的な利用は農業分野において特に有望な情報収集手段の一つになりうる。

## 1.2 リモートセンシングの概要

リモートセンシングはこれまで様々な分野で用いられてきており、農業分野においてもその利用が注目されつつある。リモートセンシングとは、離れた所から対象物に直接手を触

れないで、対象物を同定あるいは計測し、またその性質を分析する技術である（日本リモートセンシング研究会，2001）。一般的には、航空機や人工衛星などに搭載したセンサを使用し、地球表面の対象物からの電磁波を観測することにより、これら対象物の情報を得ることをいう。センサを搭載した移動体のことをプラットフォームといい、プラットフォームに航空機を利用したものは航空機リモートセンシング、衛星をプラットフォームに利用したものは衛星リモートセンシングと呼ばれる。衛星リモートセンシングには、広い地域を同時に観測する、周期が決まっている、様々な波長帯で観測可能である、などの特長があり、地表面の情報取得に広く利用されている（日本リモートセンシング学会，2011）。

リモートセンシングのセンサには、能動型、受動型の2種類がある。放射源が太陽や地表面など自然のものに由来している電磁波を観測する場合は、受動型リモートセンシングと呼ばれている。一方、衛星自身が電波を照射する場合のように、人工的に送出された電磁波を観測する場合は、能動型リモートセンシングと呼ばれている。

また、センサが観測できる電磁波にも様々なものがあり、波長の短いものから可視光、近赤外、短波長赤外、中間赤外、熱赤外、マイクロ波などがある。このうち、可視光から赤外線にかけての比較的短い波長を利用したセンサは、光学センサと呼ばれている。これまで実際に、光学センサを利用したリモートセンシングから作付作物を分類する研究や作物のバイオマス量を推定する研究は数多くなされてきた（例えば Krahwinkler and Rossmann 2013; Nordkvist *et al.*, 2012; Immitzer *et al.*, 2016; Ju *et al.*, 2010; Sarker and Nichol, 2011）。光学センサは、高い空間分解能や波長分解能を持つ（Nichol and Sarker, 2011）。そのため、作物の情報を捉えるために有用な手段と考えられる。しかし、光学センサによる観測データでは、大気による減衰が多く、特に煙霧や雲などの影響を受けやすいとされている。そのため、天候によっては作物分類や生育推定の精度が大幅に低下するおそれがあり、使用できる衛星データの数が限られてしまう可能性がある（Blaes *et al.*, 2005）。また光学センサのうち、可視光や近赤外など短波長の電磁波を受動型センサで測定する場合、暗所や夜間では微弱な放射源の下での測定になるため困難となる（McNairn *et al.*, 2002）。

### 1.3 合成開口レーダと偏波

光学センサに対して、マイクロ波を利用したセンサも多く存在している。マイクロ波を使用したセンサのうち、特に能動型センサによるものは合成開口レーダ (SAR, Synthetic Aperture Radar) と呼ばれている。能動型センサは自身で照射した電磁波を観測するため、受動型センサのように太陽光などを必要としない。そのため、昼夜を問わず安定した観測が可能である。またマイクロ波センサは、光学センサと異なり大気による減衰が少なく、大気や天候に影響をほぼ受けることなくデータを取得することが可能である (Bindlish and Barros, 2001; Mascolo *et al.*, 2019) という特長もある。したがって、SAR データを用いた作付や生育の状況に関する情報把握は、光学センサと比較しても汎用性、堅牢性が非常に高く、毎年安定して確実にデータを得られる手法であることから (石塚, 2016), 特に多時期のデータを使用する作付作物分類や生育状況推定には効果的であると考えられる。

マイクロ波は波長が 1mm から 1m の電磁波であるが、SAR において利用されている波長域には、波長の短い順に X バンド (2.5cm-3.8cm), C バンド (3.8cm-7.5cm), L バンド (15cm-60cm) などがよく使用されている。波長が長いバンドの方が植生を浸透しやすいとされており、X バンドは植物体の表面で反射し、C バンドは表面を透過して枝などで反射する (水野ほか, 2014)。そのため農業分野では、異なるバンドの衛星データを併用することにより、地表面の植生状態をより詳細に把握する効果が大きいと期待されている。

マイクロ波において、電界の振動方向は偏波と呼ばれる。偏波の種類には主に 2 種類存在し、進行方向と電界の振動方向を含む面 (偏波面) が時間的に変わらないものは直線偏波と呼ばれ、一定の速度で回転するものは円偏波と呼ばれる。また、直線偏波にも 2 種類存在し、地表面と偏波面とが水平であるものは水平偏波 (H 偏波), 地表面と偏波面とが垂直であるものは垂直偏波 (V 偏波) と呼ばれる。多偏波の SAR データでは、一般的に H 偏波, V 偏波 2 種類の偏波に対して送受信を行う。これにより、4 組の偏波データの強度を観測することが可能となる。例えば H 偏波で送信を行い、対象物での反射後に V 偏波として受信した場合、これは HV 偏波となる。このようにして、偏波における送受信の組み合わせより、HH 偏波, HV 偏波, VH 偏波, VV 偏波の 4 種類が存在することとなる。センサには、4 偏波すべてを観測するもの (4 偏波 SAR) や、HH 偏波と VV 偏波との 2 偏波のみを観測するもの (2 偏波 SAR), HH 偏波のみを観測するもの (単偏波 SAR) など多岐にわたる。これ

ら偏波におけるそれぞれの強度は後方散乱係数という数値で表され、作付作物分類や生育状況推定のための入力変数の一つとして利用が期待される。

#### 1.4 散乱成分分解

SAR データは、その偏波データを成分分解することで、後方散乱係数以外にもいくつかの値を算出することができる。この分解を行うことを散乱成分分解 (Polarimetric decomposition) といい、大きく分けて散乱モデル分解法と固有値解析法との 2 種類の方法がある。前者については、散乱のメカニズムを示す表面散乱、二回散乱、体積散乱の 3 成分、あるいはヘリックス散乱を加えた 4 成分に分解され、後者については、散乱の様相を示すエントロピ、アニソトロピ、アルファ角の 3 成分に分解される。これらの成分も後方散乱係数と同様に、分類処理における入力変数の一つとして利用が期待される。

表面散乱とは、SAR が照射したマイクロ波が地表面で奇数回 (主に 1 回) 反射する散乱過程である。地表面で発生するため、裸地状態で大きくなる傾向にある。二回散乱とは、地表面のほか壁や植生などにさらに反射し、偶数回 (主に 2 回) の反射する散乱過程である。そのため、小麦など鉛直方向の茎をもつ植生があれば大きくなる傾向にある。体積散乱は、ランダムな反射方向を有する複雑な物体の集合によって引き起こされる散乱過程である。これも植生とは関係があり、特に絡み合った枝や葉などの複雑な構造があることによって大きくなる傾向にある。ヘリックス散乱は、直線の偏波が反射後に螺旋状の円偏波に変化する散乱過程であり、人工建造物などの影響で大きくなる (山口, 2007)。

エントロピは、SAR データの散乱メカニズムである表面散乱、二回散乱、体積散乱の複雑さを示す指標であり、0 から 1 の値を取る。0 は 1 つの散乱メカニズムのみ発生していることを示し、1 は 3 つの散乱メカニズムが等しく発生していることを示している。アニソトロピも 0 から 1 の値を取り、2, 3 番目に強く発生している散乱メカニズムの相対的な力を表している。アルファ角は反射角の平均を表す数値であり、この値の大小は優位に発生している散乱メカニズムを推定する指標となる。0° から 90° の値を取り、0° 付近で表面散乱、60° 付近で体積散乱、90° 付近で二回散乱の優位性を意味する (Jiao *et al.*, 2014)。

本研究では、後方散乱係数、散乱モデル分解法による散乱成分、固有値解析法による散乱

成分を算出し、これらのデータセットのうちの一部あるいは全部を使用した。

## 1.5 SAR データ利用に関する既往研究

前項まで指摘してきたとおり、リモートセンシングを利用して作付や生育の状況を把握するためには、マイクロ波の能動型リモートセンシングである SAR データを使用する方法が、受動型センサや光学センサと比較して天候や昼夜の影響を受けにくく有利であり、多時期のデータを使用する際には効果的である。農業地域では、リモートセンシングに対する情報ニーズとして、農地利用状況の把握、作物生産性の定量化、生育診断や精密農業管理、災害実体の把握、の4点が挙げられており (井上, 2016)、SAR を利用してこれらの情報を取得することが期待できると考えられる。

### 1.5.1 作付作物の分類

農地利用状況の把握手段の開発を目指して、SAR による衛星リモートセンシングを利用して畑作地帯における圃場ごとの作付作物の分類を行う研究が行われている。

X バンド SAR データを使用した作付作物の分類では、16 時期の X バンド SAR から後方散乱係数を算出することにより、高精度な分類結果が得られることが報告されている (Sonobe *et al.*, 2014)。一方、SAR データは先述のとおり散乱成分分解により散乱成分を使用することができるが、X バンド SAR については散乱成分を作付作物の分類に適用した例は Sonobe (2019) のみと少なく、TerraSAR-X における後方散乱係数と散乱成分の併用による作付作物の分類は行われていない。また、これらの研究では 16 時期と長期にわたる複数の衛星データを使用して分類を行っているが、衛星データは観測数が多ければ多いほど取得にコストがかかるため、分類に使用する衛星データ数の削減を検討する必要がある。分類に有効な衛星データ時期を検討し、衛星データ数の削減により効率的な分類を実行した研究もみられない。出来るだけ取得数の少ない衛星データから複数の変数を用いることで、効率的かつ高精度な分類手法を検討する必要がある。

C バンド SAR データについては、近年作付作物の分類に関して多くの研究が行われている。例えば、12 時期の 4 偏波 RADARSAT-2 データから後方散乱係数を算出し、作付作物分

類を実行することで、高精度な分類結果が得られている (Skakun *et al.*, 2015)。また、RADARSAT-2 については複数の散乱成分を使用することの有効性も示されており (Jiao *et al.*, 2014; Xie *et al.*, 2019), 後方散乱係数と、散乱モデル分解法および固有値による散乱成分の計 9 変数を使用することで、作付作物分類を高精度に行える研究が報告されている (Xu *et al.*, 2014)。一方、4 偏波 SAR において散乱モデル分解法には複数の算出方法が提案されており (例えば Freeman and Durden, 1998; van Zyl *et al.*, 2011; Yamaguchi *et al.*, 2005), これらの分類の精度の比較や併用することの有用性を検討する必要があると考えられる。しかし、これら散乱成分を比較した研究は Srikanth *et al.* (2016) など限られており、複数の散乱モデル分解法による散乱成分を併用した研究や、固有値解析法による散乱成分と併用した研究は行われていない。したがって、作付作物分類により有効な変数の組み合わせを精査する必要がある。また、既往研究の多くは、X バンドと同様に長期にわたる複数の衛星データを使用して分類を行っているため、分類に使用する衛星データ数の削減を検討する必要がある。C バンドを用いたアプローチについても、出来るだけ少ない衛星データの時期数かつ複数の変数を用いて、効率的かつ高精度な分類手法を検討する必要がある。

合成開口レーダによる作付作物分類は、異なる複数のバンドを併用することで精度の向上を図れる研究が報告されており、例えば 1~2 時期の X バンドと C バンドの異なる波長のデータを組み合わせることにより、相互に補完し合うことで精度が上昇することが示されている (Zhao *et al.*, 2014)。しかし、多時期の X バンドと C バンドを併用した作付作物の分類に関する研究例は他に少なく、特に同じ時期に取得された X バンドと C バンドを併用した研究や、両波長を併用して分類に使用する際に最良の取得時期を提示した研究例はみられない。そこで、X バンドと C バンドを併用した分類についても、最良なデータ時期の組み合わせを決定し、後方散乱係数や散乱成分など複数の変数を用いて、効率的かつ高精度な分類手法を検討する必要がある。

X バンドや C バンドの他に、より波長の長い L バンドを用いた作付作物分類の研究もみられる。SAR データは基本的に有料であり、1 枚あたり数十万円することが多い。一方、L バンド SAR である ALOS-2/PALSAR-2 データは 1 枚当たり数万円と比較的安価であり (古川・古関, 2019), その有用性が確認できれば TerraSAR-X データや RADARSAT-2 データの代替としてよりコストの削減を図ることができる。また、PALSAR-2 データは既に観測した時期のデータについて、アーカイブから必要な時期と観測地域が存在すればデータを購入

することができる。このLバンドをCバンドと併用した研究例もいくつか存在しているが、1 時期での土地被覆分類においては X バンドを使用しなくてもこの 2 バンドのみで高い精度を得られることも報告されており (Qi *et al.*, 2019), L バンドが他の波長の SAR の代替として利用可能であることを示唆している。しかし、畑作地帯の作付作物分類において、多時期の RADARSAT-2 と PALSAR-2 を併用することの有効性が示されている研究はなく、各衛星から後方散乱係数だけでなく散乱成分を算出して分類に使用した例もみられないため、検討が必要である。すなわち、アーカイブで取得できる L バンド SAR データと C バンド SAR データの各分類結果の比較や、L バンド SAR データと一部の時期の C バンド SAR データを併用した分類の有効性を示すことで、低コストかつ高精度な分類を目指すためのデータ選択方法を提示する必要があるといえる。

作付作物の分類にあたっては、様々な分類手法が用いられている。その中でも Random Forests (RF) は、パターン認識の手法を用いて高い精度が期待できる機械学習アルゴリズムの 1 つであり、いくつかの研究で高精度な作付作物の分類に寄与している。RF はサポートベクターマシンなどの分類アルゴリズムと比較して高い精度を示す報告があるだけでなく (Son *et al.*, 2018; Sicre *et al.*, 2020), ノイズおよび外れ値への高い耐性があることから、複数の研究で衛星データによる高精度な分類に寄与している (Zhong *et al.*, 2014; Zhang *et al.*, 2019)。一方、RF のように、複数の決定木を使用したアルゴリズムは他にも開発されており、例えば Extremely Randomized Trees (ERT) (Geurts *et al.*, 2006) や Random Ferns (RFN) (Ozuysal *et al.*, 2010) などがある。ERT や RFN を用いた分類では、高い精度を得られた研究が報告されている。例えば ERT は脳腫瘍の自動検出 (Soltaninejad *et al.*, 2017), RFN はチームスポーツにおける選手の自動検出 (Parisot and Vleeschouwer, 2017) に利用されており、高い精度が得られている。リモートセンシングの分野でも、ERT は光学センサによる作付作物分類で有用な結果を示した報告がなされている (Kim *et al.*, 2019)。しかし、SAR データによる作付作物分類で ERT や RFN が使用された例はなく、RF との精度面の違いが不明である。

## 1.5.2 生育状況の推定

農業地域におけるリモートセンシングに対する情報ニーズとして、農地利用状況の把握の他に作物生産性の定量化が挙げられている (井上, 2016)。畑作地帯のみならず稲作地帯でも作物生育状況を定量的に推定することは重要であり、例えば圃場ごとに水稻の生育状

況を把握することは、農作業の適切な管理、収穫適期の予測、品質や収量の予測のための判断基準として利用できる。このうち特に収穫適期の予測に関しては、これまで出穂日以降の積算気温による方法で判断されてきたが、地域全体ですべての圃場の出穂日を調査することは時間的、労働的に非常に困難である (境谷・井上, 2013)。一方、出穂日がわからなくても、ある時点での穂含水率から収穫適期までの日数を予測できるため (佐々木・伊藤, 2000)、穂含水率を推定することで圃場間での水稻の収穫適期の早晩を把握することができる。この情報を用いることで、大型機械の共同利用が行われている集団営農では、各圃場の穂含水率は適切な順番で収穫を実行するための定量的な指標となり、有益であると考えられる。

既往の研究には、光学センサを利用した水稻の生育状況を推定するものが多い。例えば、Landsat データから算出された NDVI などの変数を用いることにより、水稻の草高や被覆率が推定されている (Oguro *et al.*, 2003; Yaghouti *et al.*, 2019)。一方、SAR は地表面の特性に対する強い感度を有しているため、SAR を用いた手法は生育状況の推定に対し信頼できる値を定常的に提供できると考えられる。SAR を利用した生育状況の推定も研究例は多く、X バンド SAR を使用した水稻の播種日の推定 (Mascolo *et al.*, 2019) や、水稻のフェノロジーの推定 (Kucuk *et al.*, 2016) などが挙げられる。C バンド SAR については、草高やバイオマス量を推定する研究が多く報告されている (例えば Ndikumana *et al.*, 2019; Sharifi and Hosseingholizadeh, 2019)。また、光学センサと C バンドを併用することにより、水稻の生育状態をより高精度で把握できることも示されている (Ramadhani *et al.*, 2020)。すなわち、水稻の生育状態を C バンド SAR から推定することや、光学センサを用いて精度の向上を図れることの有効性が示されている。一方、穂含水率を推定する研究として、X バンド SAR を用いた小麦の穂含水率の推定に関する研究例が報告されている (Sonobe *et al.*, 2014c)。しかし、C バンド SAR を用いた穂含水率を推定する研究や、C バンド SAR と光学センサを併用した穂含水率を推定する研究、水稻の穂含水率に注目した研究例はみられない。先述のとおり、穂含水率を推定することで収穫適期の予測ができると考えられるため、穂含水率の推定は極めて有用な情報となる。また、SAR データは様々な入射角で観測したものを使用することができるが、水稻の生育状態を推定するにあたり、この入射角の影響を比較検討した研究は、C バンド SAR の入射角ごとのバイオマス量の推定精度を比較した報告 (Valcarce *et al.*, 2019) などわずかな例しかなく、特に穂含水率を推定する研究例はみられない。くわえて、これらの研究の多くは海外での研究事例であるが、現在の北海道では、乾田直播栽培と移植

栽培の 2 種類の栽培形態で水稻を栽培している。このような栽培形態に着目した水稻の生育状態を推定する研究例もみられない。

## 1.6 本研究の目的

これまで述べたことを踏まえ、本研究では、主に SAR により取得された衛星データを利用して、作付作物の分類および生育状況の推定を高精度かつ効率的に行い、農業分野における有望な情報収集手段としての適用可能性を評価することを目的とした。作付作物の分類は畑作地帯における主要な 6 作物で行い、生育状況の推定は稲作地帯における水稻で行い、それぞれの適用可能性を評価した。

今回使用したデータは 4 種類の衛星から取得されたものであり、C バンドの SAR を搭載した RADARSAT-2、X バンドの SAR を搭載した TerraSAR-X、L バンドの SAR を搭載した ALOS-2/PALSAR-2、および高解像度の光学センサである Geoeye-1 が観測したものである。

第 1 に、X バンドの SAR を搭載した TerraSAR-X の 2 偏波データを用いた作付作物の分類を行った (第 2 章)。本研究では先行研究 (Sonobe *et al.*, 2014a) を参考として、パターン認識の手法を用いて高い精度が期待できる機械学習アルゴリズムの Random Forests (Breiman, 2001) を使用し、教師付き分類を実行した。また、分類を行うにあたり、SAR データから算出される主要な数値である後方散乱係数に加え、散乱成分分解 (Cloude, 2007) によって算出される散乱成分も変数として使用した。本研究ではまず、5 時期分の衛星データとそれぞれから得られる 4 つの要素をすべて用いて作付作物の分類を行い、精度を評価した。一方で、使用する時期数が少ないほど、衛星データの取得に費やすコストをより下げることができる。また、要素数も不要なものを省略した方が簡単に分類できる。そこで、これら時期数や要素数を削減した分類でも高い分類の精度を維持することが可能かどうかに関しても検討した。

第 2 に、C バンドの SAR を搭載した RADARSAT-2 の 4 偏波データに X バンドの SAR を搭載した TerraSAR-X の 2 偏波データを併用して、作付作物の分類を行った (第 3 章)。ここでも Random Forests を使用して教師付き分類を行った。RADARSAT-2 の 4 偏波データから

は、主要な数値の後方散乱係数だけでなく、各種の散乱成分分解手法から算出される散乱成分も作付作物の分類処理の変数として使用した。後方散乱係数と各種散乱成分の最適な組み合わせに着目した研究事例はあまりみられないが、本研究では作付作物分類において高い精度につながる最適な変数を検討し、これらの変数のうちどの変数が重要であり、どのような作物を分類するために有効であるかを考察した。なかでも、散乱成分に関しては、散乱モデル分解法として複数の手法が提唱されており、これらの手法が分類に寄与する差異についてまとめた。また、時期数を削減したうえで今一度分類を実施し、少ない時期数の衛星データで効率的に分類を行えるかについても併せて検討した。これらの結果をもとに、畑作地帯における作物分類を行うために必要な観測時期や変数について、他地域にも適用できるよう、その特徴をまとめた。

第3に、RADARSAT-2とTerraSAR-Xを使用した分類について、各種の機械学習アルゴリズムを用いた高精度化や効率化に関する検討を行った(第4章)。分類手法として、先述の機械学習アルゴリズムによる教師付き分類のRandom Forestsのほか、複数の決定木を使用したアルゴリズムが他にも開発されている。本研究では、Random Forestsのほかに、Extremely Randomized Trees (Geurts *et al.*, 2006) および Random Ferns (Ozuysal *et al.*, 2010) を使用し、精度の比較を行った。また、教師データの比率や衛星データの時期数を変更し、SARによる作付作物分類に適した分類アルゴリズムを検討した。CバンドとXバンドの各SARデータで作付作物分類を実行し、各条件下で分類の精度を比較した。とくに、同じ時期でCバンドとXバンドの比較を行い、それぞれの精度や特徴を比較した。その結果を踏まえ、使用する衛星データの利用コストを削減するために、各バンドのすべての時期のSARデータを使用せずに、同等の精度が得られる手法の提示を行った。

第4に、CバンドのSARを搭載したRADARSAT-2の4偏波データおよびLバンドのSARを搭載したALOS-2/PALSAR-2の単偏波データ、4偏波データを併用して、作付作物の分類を行った(第5章)。分類手法にはRandom Forestsを使用した。本研究では、RADARSAT-2やPALSAR-2から得られたデータを用いて、様々な条件下でそれぞれの衛星データを用いた分類における精度の比較や、両者を併用した分類における精度の比較を行った。PALSAR-2はその他のSARデータより取得コストが低い(古川・古関, 2019)。にもかかわらず、既往研究には、CバンドとLバンドを併用して衛星データ数の削減を検討した分類の研究がみられないため、本研究では特にRADARSAT-2をPALSAR-2に置換することに着目し、コ

ストをできるだけ削減した上で効率的に分類を行えるかについても併せて検討した。

第 5 に、C バンドの SAR を搭載した RADARSAT-2 の 4 偏波データおよび高解像度光学センサである Geosyde-1 データを用いて、稲作地帯における水稻の生育状況の推定を行った (第 6 章)。とくに本研究では、水稻の収穫適期との関係が強い穂含水率を推定することに焦点を当てた。モデルの作成には重回帰分析を使用し、ステップワイズ法による変数減少法から有効な変数を推定した。まず、RADARSAT-2 データによるモデルの作成を行った。モデルの作成にあたっては、衛星の進行方向や栽培形態が精度に影響を及ぼすと考え、全データを使用したモデルの他に、衛星の進行方向ごとのモデルや栽培形態ごとのモデルを作成した。また、Geosyde-1 データによるモデルや、RADARSAT-2 データと Geosyde-1 データを併用したモデルを作成した。その結果から、光学センサによるモデルの利点、欠点や、両者を併用することの有用性について検討を行った。最後に、RADARSAT-2 データから作成したモデルについて、翌年のデータを用いてモデルの検証を行った。その結果から、過去の年のデータから作成されたモデルが翌年に適用可能であるか検討を行った。

第 6 に、C バンドの SAR を搭載した RADARSAT-2 の 4 偏波データとダミー変数を用いて、穂含水率推定モデルの高精度化を検討した (第 7 章)。第 5 で述べたモデルでは、衛星の進行方向や栽培形態ごとに別なモデルを作成したが、1 つのモデルに統一することや、両方の情報をモデルに当てはめることによって、更なる高精度化を図ることを目的として、ダミー変数を使用した。ダミー変数は 0 と 1 で表される変数であり、SAR の変数のように数字で表すことのできる情報以外の情報を数字に変換したものである。本研究では、衛星の進行方向、衛星の入射角、水稻の栽培形態の 3 項目をダミー変数に変換し、モデル作成に含めることで、精度の向上を目指した。モデルの作成には重回帰分析を使用し、ステップワイズ法による変数減少法から有効な変数を推定した。また、本研究では交差検証を行い、作成したモデルの妥当性についても検証を行った。

以上、第 1 から 6 までの内容をまとめると、表 1-1 のようになる。

表 1-1 本論文の構成

章	内容	衛星データ	変数	分類・回帰手法
1) 第2章		Xバンド		RF
2) 第3章	作付作物分類 (畑作地帯)	Cバンド・Xバンド	後方散乱係数・ 散乱成分	RF
3) 第4章		Cバンド・Xバンド		RF・ERT・RFN
4) 第5章		Cバンド・Lバンド		RF
5) 第6章	生育状況推定 (稲作地帯)	Cバンド・ 光学センサ	後方散乱係数・ 散乱成分・NDVI	重回帰分析
6) 第7章		Cバンド	後方散乱係数・ 散乱成分・ダミー変数	重回帰分析

## 2 Xバンド SAR データを使用した機械学習アルゴリズムによる作付作物の分類

### 2.1 背景と目的

わが国では近年、国や地方自治体における業務合理化の一環として、統計情報の取得に費やすコストの削減が進められている。農業分野でも、現在のような作付状況の現地調査は行われにくくなるおそれがあり、現状を正確に捉えることが次第に困難になると考えられる。また農家レベルでも、農地の賃貸借の進展や法人化に伴う集団化などにより、圃場の特性を熟知して管理を行うことが難しくなっている（志賀，2010）。そのため、コストの削減を進める一方で、現在のような統計情報の取得も進めていく必要があり、これらの両立が日本の農業地域における課題の一つとなっている。

農業地域における統計情報のうち最も重要な役割を果たす項目として、作付状況が挙げられる。しかし、現在人の手によって行われているような現地調査によって、圃場ごとの作付作物の分類などの詳細を継続的かつ長期的、広域的に把握していくことが、今後は困難となることが想定される。そのため、現地調査に代わる手法の確立は、国内外問わず必要不可欠になると考えられる。そのような手法にはいくつか考えられるが、その中でもリモートセンシングの利用は、農業分野において有望な情報収集手段の一つになりうる。リモートセンシングに対する農業分野での情報ニーズとして、農地利用状況の把握、作物生産性の定量化、生育・診断や精密農業管理、災害実体の把握の4点が挙げられており（井上，2016）、このうち農地利用状況の把握をリモートセンシングで行うことは、作目の多い畑作地帯では特に重要であるといえる。

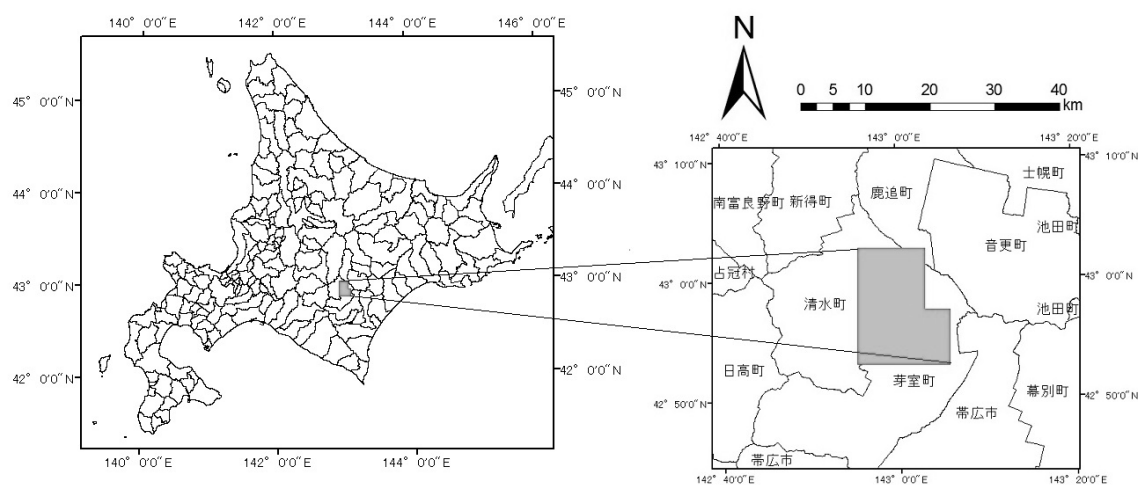
そこで本章では、先行研究（Sonobe *et al.*, 2014a）と同様に、パターン認識の手法を用いて高い精度が期待できる機械学習アルゴリズムの Random Forests（Breiman, 2001）を使用し、衛星データにおける作付作物分類の可能性について検討を行った。今回使用した衛星データは、Xバンドの合成開口レーダ（SAR）を搭載した TerraSAR-X が観測したものである。また、分類を行うにあたり、SAR データから算出される主要な数値である後方散乱係数だけでなく、散乱成分分解（Cloude, 2007）によって算出される散乱成分も変数として使用した。

本章ではまず、最高の精度で分類ができるよう、Random Forests におけるパラメータのチューニングを行った。次に、この結果を用いて、5 時期分の衛星データとそれぞれから得られる 4 つの要素をすべて用いて作付作物の分類を行い、精度を評価した。また、使用する時期数は少なければ少ないほど、衛星データの取得に費やすコストをより下げることができる。また、要素数も不要なものは省略した方が簡単に分類できる。したがって、これらの時期数や要素数を削減したうえで分類することにより、高い分類の精度を維持することが可能であるかどうかについても検討した。また最後に、これらの結果をもとにして、精度の高いデータ数で作付状況を表すマップを作成した。

## 2.2 研究方法

### 2.2.1 対象地域および対象作物

研究対象地域は、北海道十勝管内の西部に位置する河西郡芽室町、上川郡清水町の畑作地帯 (東経 142° 55' 12" ~ 143° 05' 51", 北緯 42° 52' 48" ~ 43° 02' 42") とした (図 2-1)。約 15km 西には標高 1,700m を超える急峻な日高山脈が位置しているが、対象地域内は十勝平野であり、標高は 80m~260m である。標高の高い地域は平地から連続するなだらかな斜面上に位



「国土数値情報 (行政区画データ) (国土交通省) <<https://nlftp.mlit.go.jp/ksj/index.html>> を加工して作成

図 2-1 対象地域

置しており、全体としては平坦な地形が広がっている。この地域の年平均気温は 6.1°Cで、夏季は高温になることもあるが、概ね冷涼であり、冬季の気温は氷点下となる。年平均降水量は 958mm であり、一年を通して降水がある。夏季は特に降水量が多く、冬季には積雪があり根雪となる。

この地域では大型農業機械がよく利用され、大規模農業が行われており、平均面積 2ha の大区画圃場が多く、輪作体系をとっていることから多様な種類の作物が栽培されている。そのため、衛星データを利用した作付作物の分類を行うことに適した地域であるといえる。対象作物は、この地域で広く栽培されている小麦、豆類、馬鈴薯、甜菜、牧草、トウモロコシの 6 種類とし、全 5,057 圃場を (総面積 10,119ha) 解析対象とした。なお、小麦には春播きと秋播きがあり、豆類には小豆、大豆、菜豆を含み、牧草にはチモシー、オーチャードを、トウモロコシにはスイートコーン、デントコーンをそれぞれ含んでいる。各作物の栽培暦を 図 2-2 に示す。

## 2.2.2 使用データおよび処理方法

衛星データは、X バンドのセンサを搭載した TerraSAR-X (TanDEM-X を含む) の 2 偏波 (HH および VV 偏波) データを使用した。これらは Strip Map モードによってアセンディング軌道より 11 日間隔で観測されたものを使用した。観測時の入射角は 41.7°~42.9°であり、ピクセルスペーシングは 2.75m とした。表 2-1 に、衛星データの観測時期を示す。以降、それぞれの観測時期を早い方から 1 期~5 期と記述する。なお、この TerraSAR-X データの処理レベルは、Single Look Slant Range Complex (SSC) であったため、ジオコーディングされ

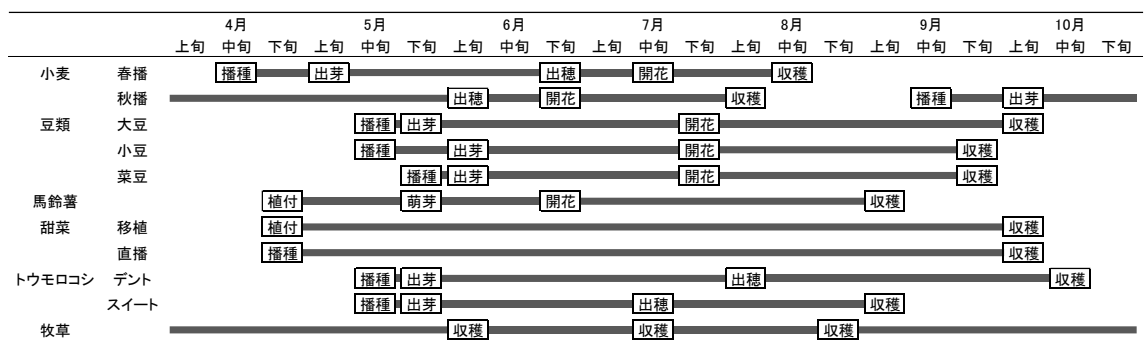


図 2-2 各作物の栽培暦

表 2-1 衛星データの観測時期

時期	観測日
1期	2013年06月17日
2期	2013年06月28日
3期	2013年07月09日
4期	2013年07月20日
5期	2013年07月31日

ていないものである。

これらの衛星データから、プロダクト仕様書 (Airbus Defence and Space, 2015) を基にして、ラジオメトリック較正を行い、後方散乱係数の算出を行った。後方散乱係数には sigma nought ( $\sigma^0$ ) を採用した。以下の式 (2-1)、式 (2-2) で HH 偏波および VV 偏波それぞれの後方散乱係数 ( $\sigma_{HH}^0$ ,  $\sigma_{VV}^0$ ) を算出した。

$$\sigma^0 = 20 \log_{10} DN + 10 \log_{10} CF + 10 \log_{10} (\sin \theta_{loc}) \quad (2-1)$$

$$DN = \sqrt{q^2 + i^2} \quad (2-2)$$

ここで、式 (2-1) の  $\sigma^0$  は後方散乱係数 (dB 値)、 $\theta_{loc}$  は衛星の入射角である。DN は衛星データのピクセル値で、センサが受信した電磁波の強さに応じたデジタル値を示しており、式 (2-2) から処理される。なお式 (2-2) の  $q$  は実部、 $i$  は虚部である。CF は校正係数であり、入射角と偏波によって異なる値をとる。

散乱成分分解 (Cloude, 2007) により、エントロピ ( $H$ ) およびアルファ角 ( $\alpha$ ) を算出した。なおエントロピとは、SAR データにおける散乱の複雑さを示す指標であり、0 から 1 の値を取る。アルファ角は、反射角の平均を表す数値であり、0° から 90° の値を取る。

以降では、後方散乱係数や各散乱成分をデータセットと呼び、各データセットの中の各偏波の後方散乱係数や散乱成分を要素と呼ぶ。以上の 4 要素 ( $\sigma_{HH}^0$ ,  $\sigma_{VV}^0$ ,  $H$ ,  $\alpha$ ) を、5 時期すべてのデータから算出した (計 20 変数)。

### 2.2.3 解析方法

分類を実行するにあたり、対象地域における圃場境界 GIS データを十勝農業共済組合よ

り提供を受けた。このデータは、各圃場の位置や区画がポリゴンとして示されており、各圃場のポリゴンに作付作物の種類など実測された属性情報が付加された GIS データである。本章ではこのデータを分類における教師データやテストデータとして使用し、圃場ポリゴン単位での分類を実行することとした。

まず、このデータのうち本章で対象とする 6 種類の作物の圃場ポリゴンのみ取り出し、それぞれに圃場 ID を与えた。また、ポリゴン内の衛星データを抽出する際、それぞれのポリゴンにおいて道路や防風林などのピクセルの混入や、マイクロ波散乱の影響が生じることを防ぎ、圃場内のデータを確実に抽出するため、圃場ポリゴンの内側に 10m のバッファを生成し、その領域を除いた。このようにして、衛星データ抽出領域の圃場ポリゴンデータを作成した。この処理には、ESRI 社製ソフトウェア ArcGIS 10.2.2 を使用した。次に、先行研究 (Sonobe *et al.*, 2014b; Sonobe *et al.*, 2018) と同様に、算出した全 20 変数について、圃場ごとの平均値を算出した。以上の方法で、衛星データ抽出領域の圃場ポリゴンデータを加工し、得られた値をその圃場の変数とした。この処理には、Hexagon Geospatial 社製ソフトウェア ERDAS imagine 2014 を使用した。

得られた全 5057 圃場分の圃場データについて、実利用において教師データを無作為に抽出することを想定し、本章でも先行研究 (Hartfield *et al.*, 2013) に倣い、ホールドアウト法により 20%の圃場を無作為に抽出した。残りの 80%の圃場はテストデータとし、精度の評価に使用した。テストデータおよび教師データの圃場数は、表 2-2 に示したとおりである。このテストデータについて教師付き分類を行い、その精度を評価した。分類処理において、本章では機械学習アルゴリズムの中でも定評のある Random Forests (以下、RF) を用いた。ま

表 2-2 解析に使用した圃場数

作物	圃場数	
	教師データ	テストデータ
豆類	187	837
甜菜	126	490
牧草	128	468
トウモロコシ	113	479
馬鈴薯	140	564
小麦	305	1,220
合計	999	4,058

た、本アルゴリズムを用いるにあたり、フリー統計ソフトである R (R Core Team, 2016) および、randomForest パッケージ (Liaw and Wiener, 2002) を活用した。

RF では、まず教師データに対してブートストラップサンプリング (Bootstrap Sampling) を行う。ブートストラップとは、重複を許した無作為な抽出のことであり、これによりいくつかのサンプルを作成する。これらのサンプルのうち 3 分の 2 を使用して、分類モデル (決定木) を作成する。ブートストラップサンプルのうち、残りの 3 分の 1 は OOB (out-of-bag) サンプルと呼ばれ、作成した決定木のテストを行う。これらを用いることで、分類モデルにおけるエラー率の評価を実施することができる。このエラー率の評価を行うことにより、作成した決定木の数 ( $n$ ) が十分であるかどうかを評価することができる。

なお分類の場合、サンプリングに用いる変数の数は、すべての変数の数の平方根を用いることが勧められている (Breiman, 2001)。そのため、本章でもそれに倣い、各決定木において用いる変数の数を 4 とした。また、RF ではジニ係数の減少量 (Mean Decrease in Gini Coefficient) を計算する。ジニ係数の減少量は、高い変数ほど分類に有効であることを示しているため、変数の重要度を推定することが可能となる。本章においても、ジニ係数の減少量を参考にしてそれぞれの変数の重要性を検討することとした。

分類の精度を評価するために使用した指標は、全体精度 (Overall Accuracy) (以下, OA) およびカッパ係数 ( $\kappa$ ) の 2 種類である。 $\kappa$  は、OA から分類が偶然的中する率を除いた精度評価指標を示すもので、その値が 0.6 以上であればある程度高い一致、0.8 以上であれば高い一致をしているといえる (Cohen, 1960)。その他、作物ごとの分類の精度の評価に、プロデューサーズ精度 (Producer's Accuracy) (以下, PA) およびユーザーズ精度 (User's Accuracy) (以下, UA) の 2 種類も使用した。PA は作物ごとの分類の精度を表しており、ある作物がその作物に正しく分類されている割合である。UA は分類結果の精度を表しており、ある作物に分類された結果のうちその作物である割合である。また、2 つの手法間における  $\kappa$  の差の有意性について、検定統計量  $Z$  (Congalton and Green, 2008) を使用して評価した。 $Z$  値が 1.96 を超えれば、5%の有意水準において、2 つの分類結果の  $\kappa$  に有意な差があると判断することができる。

## 2.2.4 解析の流れ

本章を進めるにあたって、まず OOB サンプルのエラー率から、決定木の数の検討を行う

た。次にこの決定木の数で、5 時期 (2013 年 6 月 17 日, 6 月 28 日, 7 月 9 日, 7 月 20 日, 7 月 31 日), 4 要素 ( $\sigma_{HH}^0, \sigma_{VV}^0, H, \alpha$ ) の全 20 変数を用いた分類を行った。この結果について、どの程度の精度が得られるかを評価した。最後に、これらの要素数や時期数を削減して使用した場合についても、再度 RF を適用した。削減した分類の結果から、時期数や要素数を減少させたことによって、十分な精度を維持した分類を行うことが可能かどうかを検証した。

## 2.3 結果と考察

### 2.3.1 決定木の数の検討結果

OOB サンプルを用いて、決定木の数について検討を行ったところ、決定木の数とエラー率との関係は図 2-3 のとおりとなった。この結果から、決定木の数がおよそ 50 本以上で OOB

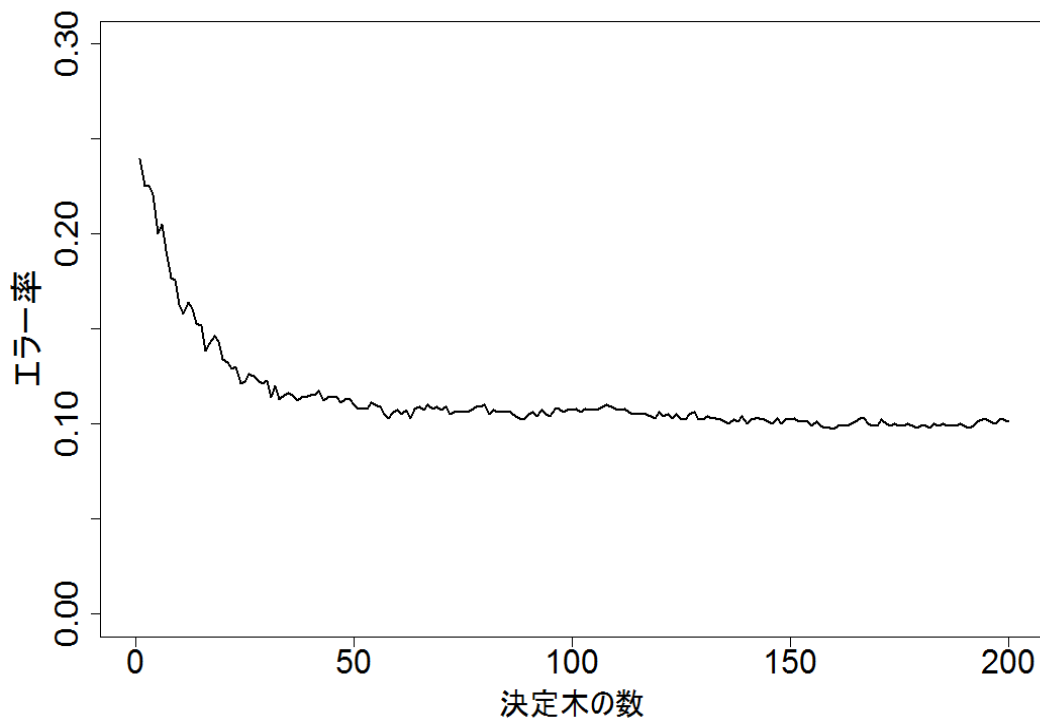


図 2-3 決定木の数による OOB サンプルのエラー率の変化

サンプルのエラー率が低く安定していることが判明した。分類において、決定木の数が多いことに関しては問題ないとされているため (Breiman, 2001), 決定木の数をそれより大きい  $n=200$  と設定し, 残り 80% のデータを分類して精度を評価した。

### 2.3.2 全変数利用時の分類の精度

5 時期 4 要素すべての変数を使用した際の判別効率表は, 表 2-3 に示すとおりとなった。すべての時期, 要素を用いた場合において, OA は 0.895 であり,  $\kappa$  は 0.869 ( $>0.8$ ) であった。以上の結果から, RF を TerraSAR-X データに適用することによって, 0.9 に近い精度で分類を行うことができた。

また, 分類モデル作成時に算出されたジニ係数の減少量 (Mean Decrease Gini) とそれぞれの変数との関係は, 図 2-4 のとおりとなった。ここに示した各パラメータのジニ係数の減少量を要素別に確認していくと, H や  $\alpha$  と比較して HH 偏波や VV 偏波の各後方散乱係数の方が特に大きく本結果に寄与していることが判明した。さらに時期別に見ると, 7 月 9 日のデータの重要度が最も高く, 次いでその前後である 6 月 28 日や 7 月 20 日のデータの重要度が高いことが示された。

これらの原因として, 7 月初旬やその前後の時期は作物種ごとの生育状況の違いが大きくなっていったことが挙げられる。そのため, それぞれの作物種の群落構造の差も明瞭となり, 後方散乱係数の変化がよく出ていたと考えられる。このことは, 7 月初旬頃に最も生育度合

表 2-3 全変数利用時の分類の判別効率表

		参照データ							
		豆類	甜菜	牧草	トウモロコシ	馬鈴薯	小麦	計	UA
分類 結果	豆類	753	15	2	41	46	6	863	0.873
	甜菜	11	442	1	12	4	0	470	0.940
	牧草	13	4	423	28	9	46	523	0.809
	トウモロコシ	29	26	15	375	8	13	466	0.805
	馬鈴薯	30	2	2	12	490	8	544	0.901
	小麦	1	1	25	11	7	1147	1,192	0.962
	計	837	490	468	479	564	1,220	4,058	
	PA	0.900	0.902	0.904	0.783	0.869	0.940		
OA	0.895	$\kappa$	0.869						

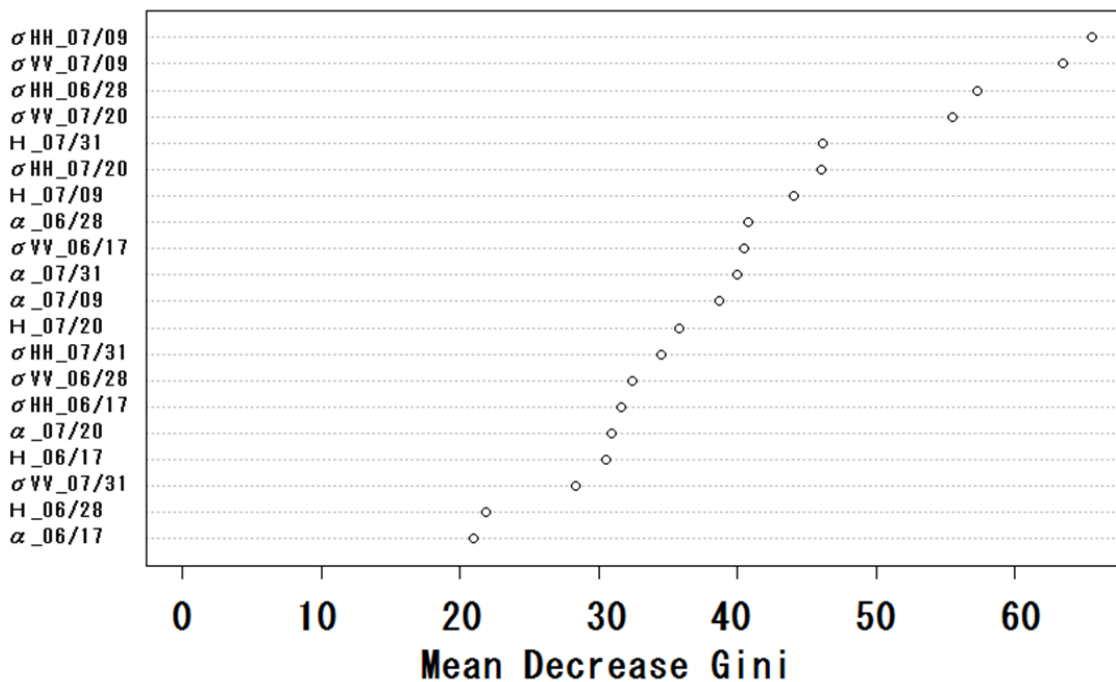


図 2-4 ジニ係数の減少量と変数との関係

が大きい秋播き小麦において、UA, PA とともに群を抜いて精度が大きくなっていたことから裏付けられる。

### 2.3.3 要素数別の分類の精度

本章において用いた要素は、 $\sigma_{HH}^0$ ,  $\sigma_{VV}^0$ , H,  $\alpha$  の 4 種類であったが、これらの要素を削減して分類を行い、その精度評価を行った。いずれか 1 要素を用いた場合、後方散乱係数の 2 要素を用いた場合 ( $\sigma_{HH}^0 + \sigma_{VV}^0$ )、散乱成分の 2 要素を用いた場合 (H+ $\alpha$ ) のそれぞれについて同様に RF による分類を実施した。以上の分類結果を 4 要素すべてを用いた 3.2 の結果と比較し、OA や  $\kappa$  にどれほどの差異が現れるかを評価した (表 2-4)。

表 2-4 によると、4 要素すべてを用いた場合と比較して、要素数の削減により精度は低下することが明らかとなった。とくに 1 要素しか用いなかった場合について見ると、OA は 0.730~0.778 であり、すべて用いた場合より 20 ポイント程度低く、 $\kappa$  は 0.665~0.725 であった。2 要素を用いた場合については OA,  $\kappa$  ともに向上し、OA は 0.830~0.841% となったが、 $\kappa$  は 0.8 前後であった。

表 2-4 要素数別の分類精度

		$\sigma_{HH}^0$	$\sigma_{VV}^0$	$\sigma_{HH}^0 + \sigma_{VV}^0$	$\alpha$	H	$\alpha + H$	All
PA	豆類	0.722	0.768	0.817	0.724	0.766	0.847	0.900
	甜菜	0.861	0.841	0.878	0.567	0.729	0.739	0.902
	牧草	0.808	0.774	0.850	0.684	0.598	0.838	0.904
	トウモロコシ	0.503	0.539	0.658	0.539	0.666	0.716	0.783
	馬鈴薯	0.729	0.670	0.771	0.729	0.809	0.860	0.869
	小麦	0.902	0.878	0.907	0.893	0.864	0.920	0.940
UA	豆類	0.689	0.688	0.751	0.647	0.828	0.822	0.873
	甜菜	0.894	0.866	0.898	0.584	0.704	0.767	0.940
	牧草	0.724	0.687	0.758	0.606	0.628	0.769	0.809
	トウモロコシ	0.560	0.590	0.706	0.594	0.617	0.713	0.805
	馬鈴薯	0.731	0.738	0.850	0.853	0.774	0.900	0.901
	小麦	0.921	0.915	0.933	0.907	0.860	0.941	0.962
OA		0.778	0.770	0.830	0.730	0.766	0.841	0.895
$\kappa$		0.725	0.715	0.789	0.665	0.710	0.803	0.869

つづいて、1 要素、2 要素のみを用いた場合について、すべての要素を用いた場合との間でそれぞれ Z 値を算定した (表 2-5)。この結果より、すべての要素を用いた分類結果は、1 要素、2 要素のみを用いた場合と比較して、5%の有意水準において有意な差のある高い分類の精度であることが明らかとなった。従って、要素数を削減することは精度の低下につながり、高精度な分類を維持することが困難になると判断した。

表 2-5 4 要素全てを用いた場合との Z 値

要素	Z 値
$\sigma_{HH}^0$	14.80
$\sigma_{VV}^0$	15.70
$\sigma_{HH}^0 + \sigma_{VV}^0$	8.66
$\alpha$	20.24
H	16.17
$\alpha + H$	7.26

### 2.3.4 時期数別の分類の精度

今回使用したデータは5時期であったが、これらをすべては使用せず一部の時期のみ使用した場合、OAや $\kappa$ にどれほどの差異があるのか評価した。なお、各時期数において用いた値は、最も精度の高い時期の組み合わせのものである(表2-6)。

時期数ごとの分類の精度の結果では、どの時期数に対しても小麦のUA、PAは常に高い値を示した(表2-7)。これは、小麦が他の作物に比べて生育が早く、後方散乱係数などにも差異が大きかったために、少ない時期数でも他の作物との区別が付きやすかったと考えられる。また、甜菜のUAやPAが2時期までの間で高い値を示したのも、移植栽培であるた

表 2-6 各時期数における使用したデータ

時期数	日付
1時期	7/9
2時期	6/28, 7/9
3時期	6/28, 7/9, 7/20
4時期	6/28, 7/9, 7/20, 7/31
5時期	6/17, 6/28, 7/9, 7/20, 7/31

表 2-7 時期数別の分類精度

時期数		1	2	3	4	5
PA	豆類	0.787	0.839	0.853	0.891	0.900
	甜菜	0.800	0.867	0.869	0.882	0.902
	牧草	0.694	0.842	0.878	0.895	0.904
	トウモロコシ	0.551	0.706	0.724	0.789	0.783
	馬鈴薯	0.768	0.853	0.855	0.858	0.869
	小麦	0.879	0.895	0.914	0.932	0.940
	UA	豆類	0.792	0.818	0.827	0.879
甜菜		0.798	0.928	0.932	0.927	0.940
牧草		0.649	0.720	0.758	0.788	0.809
トウモロコシ		0.620	0.741	0.727	0.786	0.805
馬鈴薯		0.738	0.841	0.878	0.888	0.901
小麦		0.878	0.936	0.953	0.959	0.962
OA		0.775	0.846	0.861	0.886	0.895
$\kappa$	0.721	0.809	0.828	0.859	0.869	

めに同様の理由のためと考えられる。

また、時期数と OA および  $\kappa$  との関係をグラフにして比較した (図 2-5)。その結果、総じて時期数の削減に応じて OA,  $\kappa$  ともに下降していくことが判明した。ただし、4 時期の OA は 0.886,  $\kappa$  は 0.859 であり、5 時期すべてを用いた場合と 5% の有意水準で有意な差はなかった ( $Z=1.22$ )。一方、3 時期と 4 時期の間の Z 値は 3.43 ( $>1.96$ ) と  $\kappa$  に有意な差があった。しかし、3 時期のデータを用いた場合の OA は 0.861,  $\kappa$  は 0.828 であり、3 時期にしてもある程度高い精度を維持していた。

2 時期まで減らした場合でも、OA は 0.846,  $\kappa$  は 0.809 となっており、まだ精度は高かった。また、3 時期との Z 値は 2.02 ( $>1.96$ ) となっており、有意な差がみられた。1 時期のみの使用では、OA は 0.775 と低く、 $\kappa$  も 0.721 と、高い一致はしなかったことが確認された。

以上の結果より、5 時期すべてのデータを用いなくても、十分な精度が得られることが示された。とくに 5 時期と 4 時期とでは精度がほとんど変わらず、3 時期や 2 時期でも  $\kappa$  は

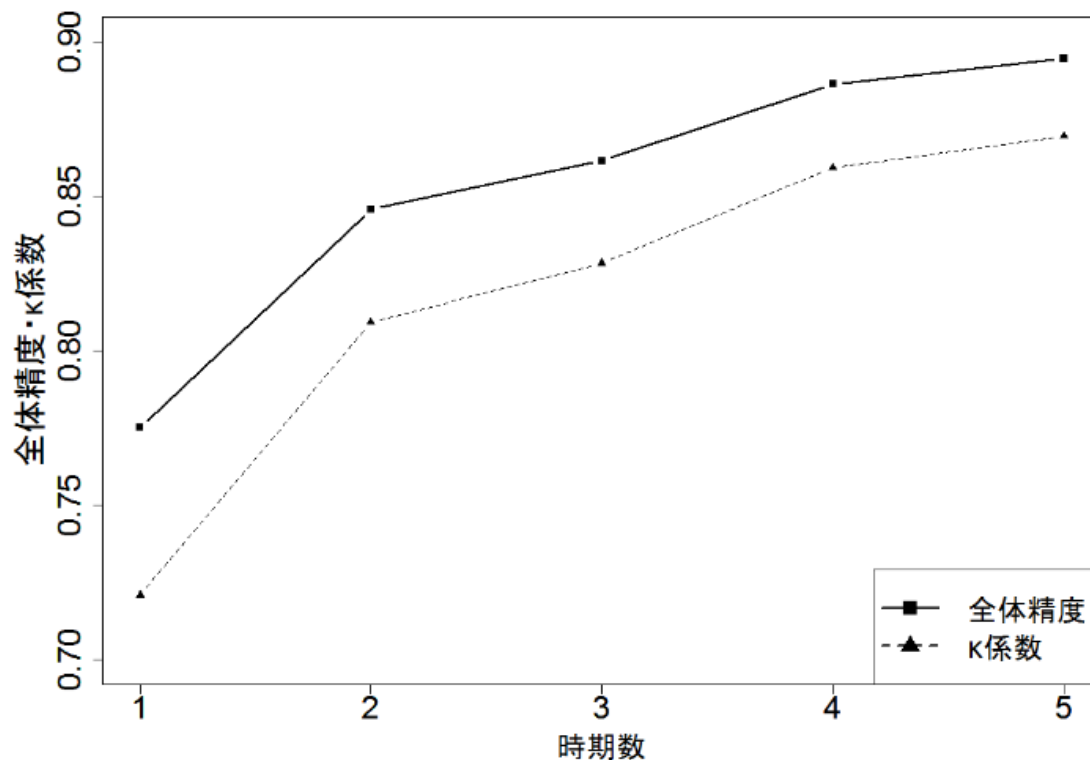


図 2-5 時期数による精度の変化

0.8 を上回る十分高い精度を得られることが確認された。そのため、2 時期以上のデータを用いれば高い精度で分類できると考えられる。また、さらに高い精度が必要な場合は、4 時期以上のデータを用いることが望ましいと考えられる。

### 2.3.5 作付マップの作成

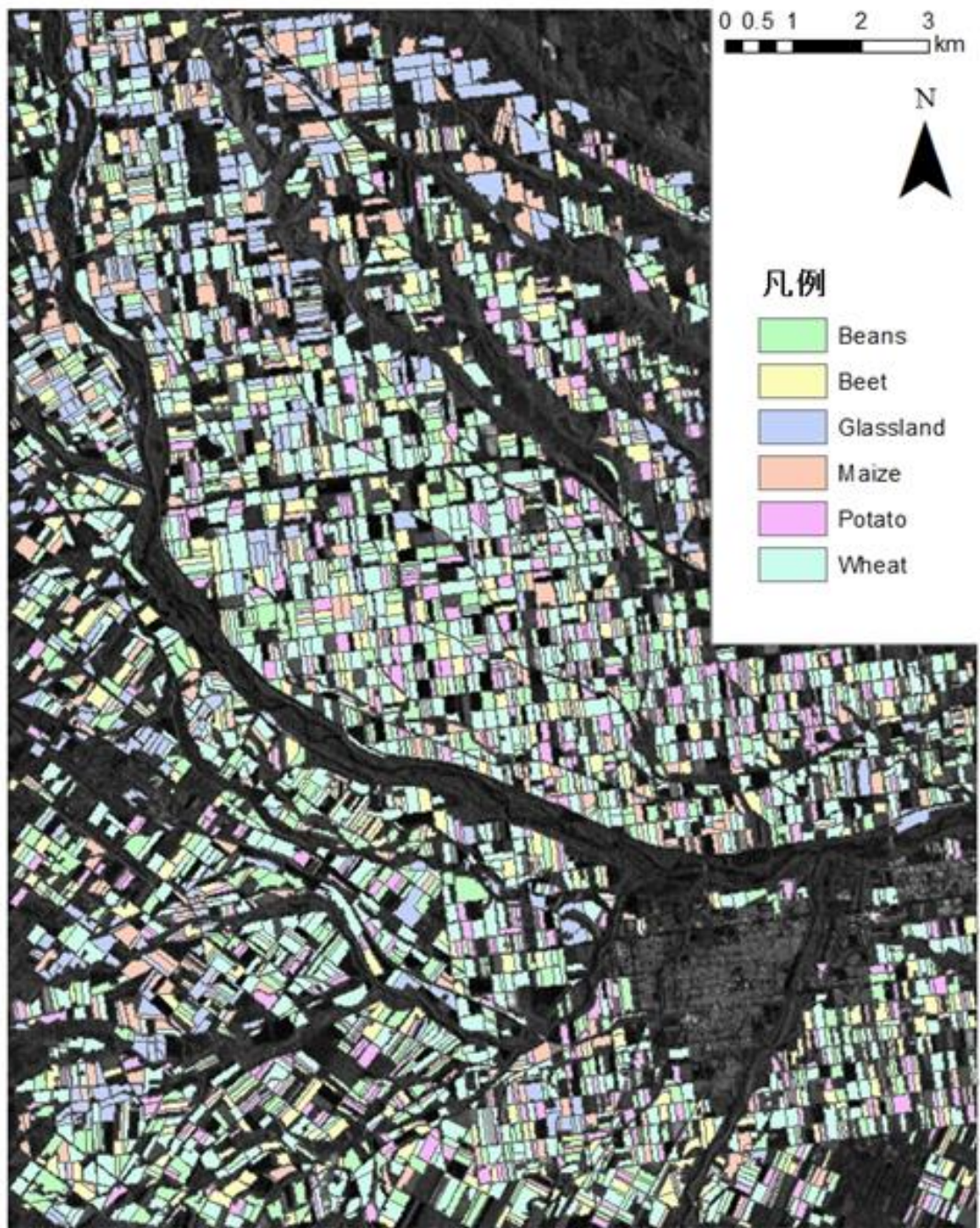
図 2-6 に、4 時期のデータを用いて作成した作付マップを示す。このように、RF によって分類した結果より、作付マップを作成することができた。また、図 2-7 ではこの作付マップにおいて誤分類された圃場を黒で示した。

誤分類された箇所は各地に点在しているが、それらの多くは数箇所の誤分類された圃場の集団や、面積の小さな圃場であることが判明した。前者の場合、倒伏や生育ムラなど作物自体に変化があり、数値が異常値となって誤分類された可能性が考えられる。後者の場合、異常値のピクセルが平均値の算出で大きく影響を与えた可能性が考えられる。

## 2.4 まとめ

本章では、TerraSAR-X の HH および VV 偏波データについて、機械学習アルゴリズムである Random Forests を使用し、作物の分類を行った。同時に、高い分類の精度を維持しつつデータ数を削減することが可能であるかどうかという視点からも検討を進めた。使用した要素は、後方散乱係数 ( $\sigma_{HH}^0$ ,  $\sigma_{VV}^0$ ) と、散乱成分分解により算出された H,  $\alpha$  の 4 要素であった。また衛星データには、2013 年 6/17~7/31 に 11 日間隔で取得した 5 時期の TerraSAR-X データを使用した。

その結果、5 時期 4 要素すべてを用いた場合、0.9 近い精度を得ることができた。要素数を削減すると精度は大きく下がり、高精度の維持は望めないことが確認された。一方、時期数を削減した場合でもある程度高い精度を維持することが可能であった。とくに 6/28 と 7/9 を含む 2 時期以上のデータ数があれば精度は十分高く、4 時期以上あれば更に高くなるため、より低コストでの作物分類が可能となると結論づけた。



© 2015 DLR, Distribution Airbus DS / Infoterra GmbH, Sub-Distribution [PASCO]

図 2-6 RF による作付マップ



© 2015 DLR, Distribution Airbus DS / Infoterra GmbH, Sub-Distribution [PASCO]

図 2-7 誤分類の生じた圃場



### 3 XバンドおよびCバンド SAR データを併用した機械学習アルゴリズムによる作物分類の高精度化・効率化に関する検討

#### 3.1 背景と目的

1章および2章で述べたように、SAR データを用いた分類においては、高精度化や少ないデータ数での効率化を図ることが求められている。先行研究では、Xバンド SAR データを使用することにより高い精度での作物分類を行うことが報告されており (Sonobe *et al.*, 2014a), 2章でも5時期のデータを使用することで0.895の全体精度が得られた。

一方、異なる波長のデータを組み合わせることにより、相互に補完し合うことで精度を向上できることも示されている (Zhao *et al.*, 2014)。

そこで、本章ではCバンド SAR データのみを使用した手法と、Cバンドに加えXバンドの SAR データを併用した手法の2種類について、畑作地帯における作付作物の分類を行い、使用するデータセットの違いによる精度を比較した。その結果から、分類に有用なデータセットの組み合わせの提示を行った。分類の手法には、2章と同様に、パターン認識の手法を用いて高い精度が期待できる機械学習アルゴリズムの1つである Random Forests を使用した。

分類を行うにあたり、SAR データから算出される主要な数値の後方散乱係数だけでなく、各種の散乱成分分解手法から算出される散乱成分も作付作物の分類処理の変数として使用した。後方散乱係数や各種散乱成分の最適な組み合わせに着目した例はあまりみられないため、本章では作付作物分類において高い精度につながる最適変数を検討し、これらの変数のうちどのような変数が重要であり、どのような作物を分類するために有効であるかを考察した。

特に4偏波 SAR データからは、散乱モデル分解法を用いた散乱成分を算出することができるため、本章では RADARSAT-2 データからこの散乱成分の算出を行った。散乱モデル分解法にはいくつかの手法が提唱されており、Freeman-Durden 法 (Freeman and Durden, 1998), van Zyl 法 (van Zyl *et al.*, 2011), Yamaguchi 法 (Yamaguchi *et al.*, 2005) の3種類の手法を用いてそれぞれの成分の比較を行った。

また、固有値解析法である Cloude-Pottier 法 (Cloude and Pottier, 1996) から算出される散

乱成分も算出し、後方散乱係数も含めたこれら 5 種類のデータセットの分類の精度の比較や、併用することによる精度の向上の検討を行った。また、TerraSAR-X データについても 2 章と同様に後方散乱係数と Cloude 法 (Cloude, 2007) による固有値解析法から算出される散乱成分を用いて、双方のバンドの衛星データを併用し精度が向上するかについて検討を行った。

衛星データは観測の時期数が多いほど、取得にコストがかかるが、分類に使用する衛星データ数の削減を検討した研究もあまりみられない。したがって、2 章と同様に、これらの時期数を削減したうえで今一度分類を実施し、少ない時期数の衛星データで効率的に分類が行えるかについても併せて検討した。これらの結果をもとに、畑作地帯における作物分類を行うために必要な時期や変数について、他地域にも適用できるよう、その特徴を簡潔にまとめた。

## 3.2 研究方法

### 3.2.1 使用データおよび処理方法

本章の対象地域および対象作物は 2 章と同じである。使用した衛星データは、C バンドのセンサを搭載した RADARSAT-2 の 4 偏波データおよび X バンドのセンサを搭載した TerraSAR-X の 2 偏波データである。RADARSAT-2 は、Fine Quad モードによりアセンディング軌道で観測されたものを使用した。入射角は  $39.2^{\circ}$  ~  $40.7^{\circ}$  (FQ20)、ピクセルスペーシングは 8.0m、処理レベルは Single Look Complex (SLC) である。また TerraSAR-X は、Strip Map モードによりアセンディング軌道で観測されたものを使用した。組み合わせた分類を行う際に衛星ごとでデータの差が生じないように、入射角は RADARSAT-2 とほぼ同等である  $41.7^{\circ}$  ~  $42.9^{\circ}$  のデータを使用した。ピクセルスペーシングは 2.75m、処理レベルは Single Look Slant Range Complex (SSC) である。取得した時期は 2015 年の 5 時期であり、表 3-1 にそれぞれの衛星データの観測日を示す。以降、それぞれの観測時期を早い方から 1 期~5 期と記述する。なお、両者の衛星データは観測された時期が近くなるようにそれぞれ選定した。

作付作物分類を実施するにあたり、ノイズ除去や座標変換のための前処理を行った。スペクトルフィルタとして Lee フィルタ (Lee *et al.*, 1999) (5×5 ピクセル) で処理し、マルチル

表 3-1 各衛星データの観測日

時期	RADARSAT-2	TerraSAR-X
1期	2015年 5月26日	2015年 5月22日
2期	2015年 6月19日	2015年 6月24日
3期	2015年 7月13日	2015年 7月16日
4期	2015年 8月 6日	2015年 8月 7日
5期	2015年 8月30日	2015年 8月29日

ック処理 (レンジ 1, アジマス 2) を行った。また, オルソ補正を行うために, 国土地理院による基盤地図情報ダウンロードサービス (国土地理院, 2016) を通じて入手した基盤地図情報数値標高モデル (10m メッシュ) およびジオイドモデル Earth Gravitational Model 2008 (NGA GEOnet Names Server, 2013) を用いた。これら一連の処理には, 欧州宇宙機関 (European Space Agency: ESA) によって頒布されている SNAP Sentinels platform application 2 (SNAP) (ESA, 2015) を使用した。

このように調整したデータを用いて, 分類に使用する変数を以下に示す方法で算出した。RADARSAT-2 データからは, 各偏波の後方散乱係数 ( $\sigma^0$ ) を (3-1) 式から算出した (MDA, 2004)。

$$\sigma^0 = 20 \log_{10}(DN/A2) + 10 \log_{10}(\sin \theta_{loc}) \quad (3-1)$$

ここで,  $\theta_{loc}$  は電波の入射角を示し, DN は衛星データのピクセル値を, A2 は入射角と偏波によって異なるスケーリングゲイン値を示している。なお, 4 偏波の衛星データにおける HV 偏波と VH 偏波は, クロストークの影響を除去した較正においてモノスタティックでの計測の場合は値が等しくなる (van Zyl, 1990) ことを考慮し, RADARSAT-2 データから算出される  $\sigma_{VH}^0$  は使用しないこととした。

$\sigma^0$  のほかに, 本章では 4 種類の散乱成分を算出した。使用した散乱成分は, 散乱モデル分解法である Freeman-Durden 法 (以下, FD) (Freeman and Durden, 1998), van Zyl 法 (以下, VZ) (van Zyl *et al.*, 2011), Yamaguchi 法 (以下, YG) (Yamaguchi *et al.*, 2005) の 3 種類の手法から算出される散乱成分と, 固有値解析法である Cloude-Pottier 法 (以下, CP) (Cloude and Pottier, 1996) から算出される散乱成分である。以降では, それぞれの  $\sigma^0$  や散乱成分をデータ

セットと呼ぶ (表 3-2)。RADARSAT-2 データからは、以上の 5 手法より計 16 要素を 5 時期すべてのデータから算出した (計 80 変数)。

TerraSAR-X データからは、2 章と同様に各偏波の $\sigma^0$  (Airbus Defence and Space, 2015) および固有値解析法である Cloude 法 (以下, CD) (Cloude, 2007) の散乱成分を算出した (2 手法, 計 4 要素)。以上の 4 要素を、5 時期すべてのデータから算出した (計 20 変数)。

一例として、6 種類の作物が作付けされた圃場群における、各時期の一部の変数を用いたカラー合成画像を図 3-1 に示す。図 3-1 より、各時期、各作物におけるそれぞれの変数の時間的、空間的差異が確認できる。

### 3.2.2 解析方法

本章でも 2 章と同様に、十勝農業共済組合より提供を受けた対象地域の圃場情報 GIS データを分類に使用した。それぞれのポリゴンにおける道路や防風林などのピクセルの混入や、マイクロ波散乱の影響が生じることを防ぎ、圃場内のデータを確実に抽出するため、圃場ポリゴンの内側にバッファを生成し、その領域を除いた。使用した RADARSAT-2 データおよび TerraSAR-X データのオルソ補正後のピクセルスペーシングがそれぞれ 8m, 2.75m であることを考慮し、圃場ポリゴンの内部面積が同じになるように、バッファは 10m に設定した。対象地域には、幅 20m 以下の狭小な圃場も数 10 か所存在したが、数は少ないため、研究対象から除外した。各作物の圃場面積別のヒストグラムを図 3-2 に示す。算出した全ての変数について、圃場ごとに平均値を抽出した。

分類処理は、2 章と同様にホールドアウト法により圃場を教師データとテストデータに分けて実行した。教師データとテストデータに用いた作物種別の圃場数を表 3-3 に示す。本章

表 3-2 算出した各データセットの要素

衛星	データセット	算出した要素			
RADARSAT-2	$\sigma^0$	HH偏波 ( $\sigma_{HH}^0$ )	HV偏波 ( $\sigma_{HV}^0$ )	VV偏波 ( $\sigma_{VV}^0$ )	
	CP	エントロピ (H)	アニソトロピ (A)	アルファ角 ( $\alpha$ )	
	FD	二回散乱 (Pd)	体積散乱 (Pv)	表面散乱 (Ps)	
	VZ	二回散乱 (Pd)	体積散乱 (Pv)	表面散乱 (Ps)	
	YG	二回散乱 (Pd)	体積散乱 (Pv)	表面散乱 (Ps)	ヘリックス散乱 (Pc)
TerraSAR-X	$\sigma^0$	HH偏波 ( $\sigma_{HH}^0$ )	VV偏波 ( $\sigma_{VV}^0$ )		
	CD	エントロピ (H)	アルファ角 ( $\alpha$ )		

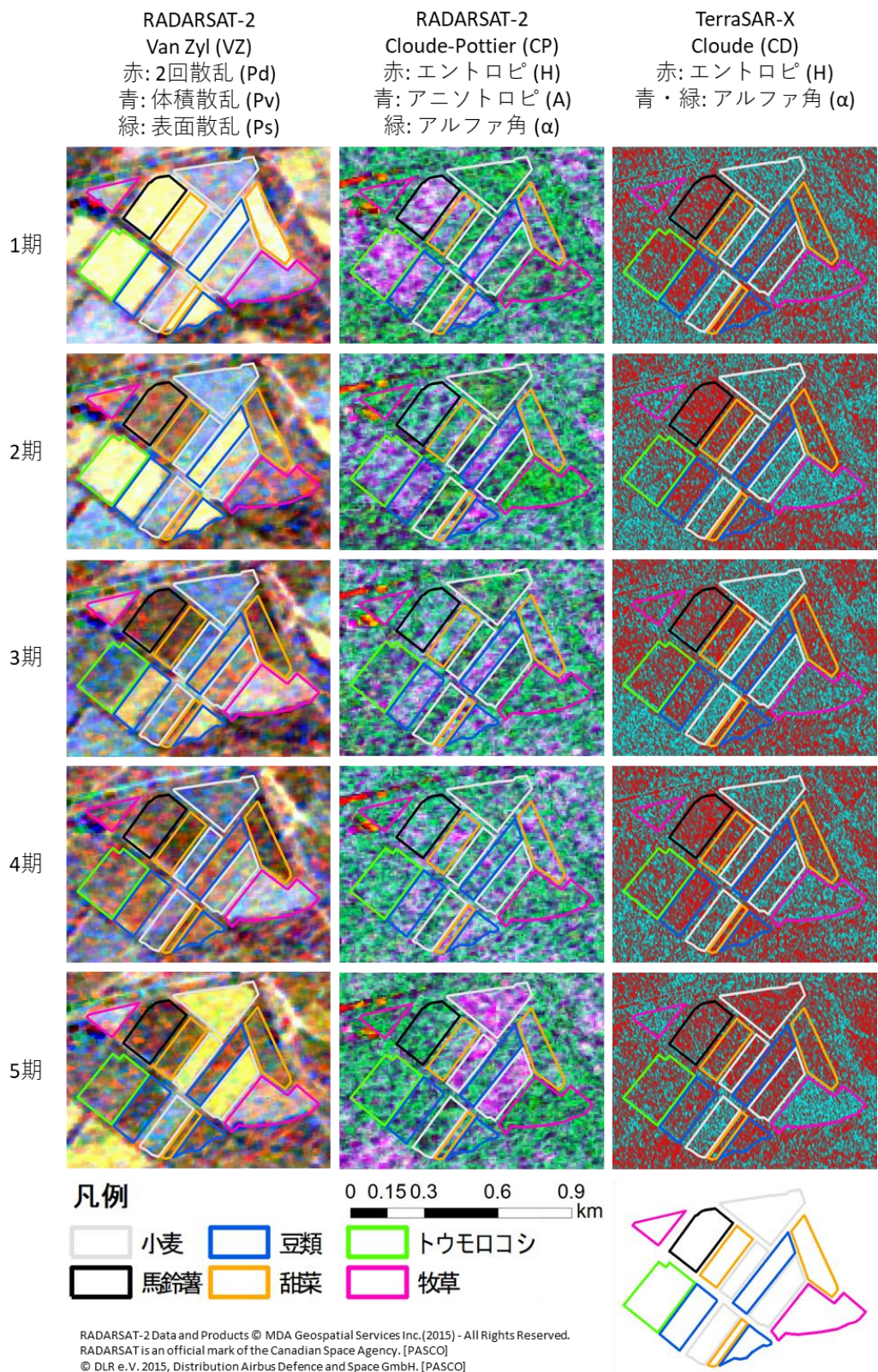


図 3-1 一部の 변수を用いた各時期のカラー合成画像の例

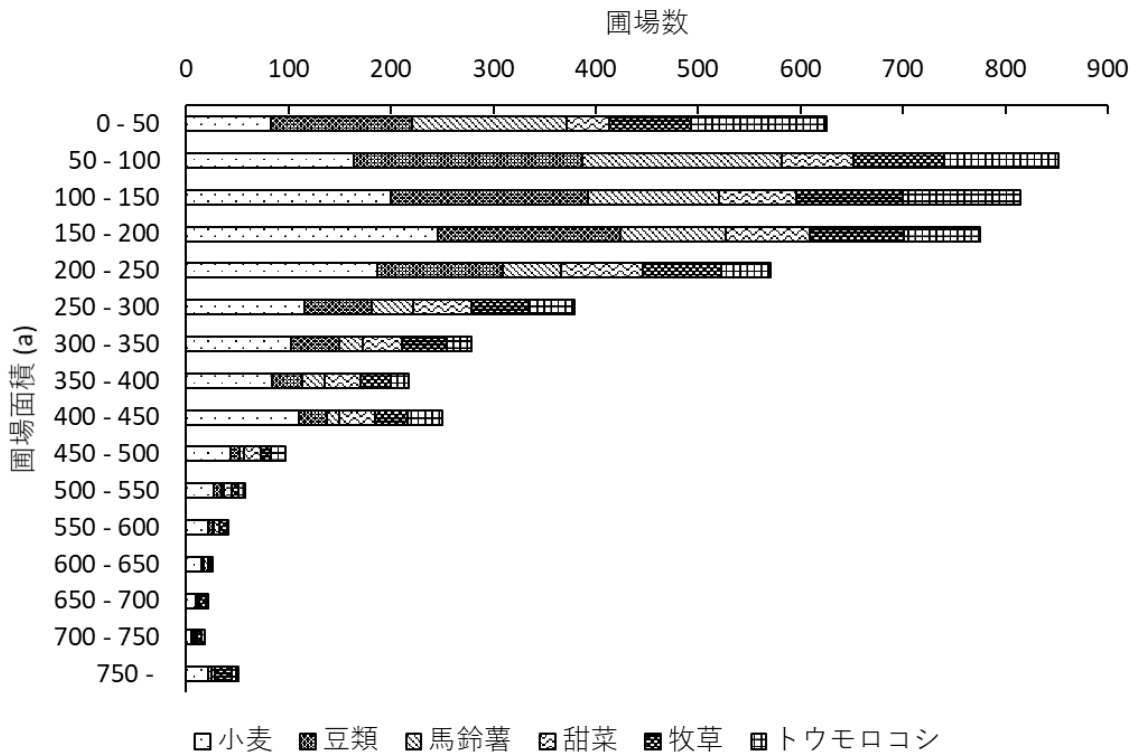


図 3-2 圃場面積別のヒストグラム

表 3-3 各圃場数の内訳

作物	圃場数		合計
	教師データ	テストデータ	
小麦	267	1,167	1,434
豆類	211	835	1,046
馬鈴薯	151	595	746
甜菜	114	444	558
牧草	137	516	653
トウモロコシ	135	504	639
合計	1,015	4,061	5,076

の対象地域には、計 5,076 の圃場があり、特に小麦 (1,434 圃場) や豆類 (1,046 圃場) が多く、他の 4 作物は 558~746 圃場であった。教師データは計 1,015 圃場、テストデータは計 4,061 圃場とした。このテストデータに対する分類結果を基に精度を評価した。

本章でも、分類には Random Forests (RF) を使用し、分類の精度の評価には全体精度 (OA)、カッパ係数 ( $\kappa$ )、プロデューサーズ精度 (PA)、ユーザーズ精度 (UA) の 2 種類も使

用した。2つの手法間における  $\kappa$  の差の有意性について、検定統計量  $Z$  を使用して評価した。

### 3.2.3 解析の流れ

本章では、以下の 1) ~ 3) で示す 3 処理で解析を進めた。

#### 1) RADARSAT-2 データを使用した分類 (2.3.2 参照)

2015 年に観測された 5 時期の RADARSAT-2 データのみを使用して、作付作物の分類を行った。まず、RADARSAT-2 データから算定に用いた 5 種のデータセット ( $\sigma^0$ , CP, FD, YG, VZ) のうち、いずれか 1 データセットのみを使用した分類により精度の比較を行った。その結果から、どのデータセットで高い精度が得られるかについて着目し、検討を行った。使用した変数の数は、YG のみ 5 時期 4 要素の計 20 変数となり、他は 5 時期 3 要素の計 15 変数となる。

次に、使用する変数を増加させて分類を行った。FD, YG, VZ の 3 データセットはすべて散乱モデル分解法であり、同一種類と考え、これらのうちいずれか 1 種類のデータセットに、 $\sigma^0$  と固有値解析法である CP を加え、算出される成分の異なる 3 種類のデータセット (後方散乱係数, 散乱モデル分解法, 固有値解析法) を組み合わせた分類を行った。これより、 $\sigma^0$ +CP+FD,  $\sigma^0$ +CP+YG,  $\sigma^0$ +CP+VZ の 3 とおりの組み合わせにおいて精度に上昇がみられるかの検討や、散乱モデル分解法の 3 データセット (FD, YG, VZ) のうちどの算定法が最も精度の上昇に寄与するかの比較を行った。なお、使用した変数の数は、 $\sigma^0$ +CP+YG のみ 5 時期 10 要素の計 50 変数となり、他は 5 時期 9 要素の計 45 変数となる。

最後に、RADARSAT-2 データにより算出される 5 種類すべてのデータセット (5 時期 16 要素, 計 80 変数) を使用して分類を行った。これにより、複数の散乱モデル分解法を組み合わせることで、精度の上昇に寄与するかの検討を行った。

#### 2) RADARSAT-2 データと TerraSAR-X データとを組み合わせた分類 (2.3.3 参照)

異なる波長のデータを組み合わせると、それらが相互に補完し合うことにより分類の精度が上昇することも示されている (Zhao *et al.*, 2014)。そこで、2015 年に観測された 5 時期の RADARSAT-2 データに、TerraSAR-X データを組み合わせることで分類を行った。本解析では、異なる波長の衛星データを併用することで、精度の上昇がみられるかを検討した。また、特にどの要素が精度を上昇させるのに効果があったかについても検討を行った。なお、

RADARSAT-2 データは 1) で最も精度の高かった変数の組み合わせを使用した。

まず、RADARSAT-2 データに TerraSAR-X の  $\sigma^0$ 、または CD のいずれかを組み合わせた分類を行った。次に、RADARSAT-2 データに TerraSAR-X の  $\sigma^0$  と CD との両方を組み合わせた分類を行った。前者は、5 時期 2 要素の計 10 変数を RADARSAT-2 データから算出された変数に加えることになり、後者は、同様に 5 時期 4 要素の計 20 変数を加えることになる。

### 3) 使用する時期数の検討 (2.3.4 参照)

衛星データは時期数が多ければ多いほど、取得にコストがかかる。そのため、時期数の削減可能性を検討するために、今一度分類を実施し直した。2 章では、TerraSAR-X の X バンドデータにおいて、4 時期のデータを使用することで 5 時期を使用した場合と有意差のない  $\kappa$  が得られた。本章では、RADARSAT-2 を用いた場合、および RADARSAT-2 と TerraSAR-X を併用した場合でも、同様に時期数を削減し、精度の維持が可能か検討した。1)、2) の結果それぞれについて、最も高い精度であったデータセットの組み合わせを使用し、5 時期すべての衛星データを使用しなくても同様の高精度な分類が可能であるかどうかを評価した。具体的には、1 時期～5 時期を使用した場合の各々について分類を行い、異なる時期数の分類間で精度にどのような差異が生じるかを検討した

## 3.3 結果と考察

### 3.3.1 現地調査結果

衛星データを取得した 5 時期について、それぞれの作物の状態を確認するため、現地調査を行った。これら 5 時期における各作物の圃場の様相は、図 3-3～3-7 に示す。



図 3-3 1 期（2015 年 5 月 25 日）の各圃場の様子（左上：小麦，右上：小豆，  
左中：甜菜，右中：馬鈴薯，左下：牧草，右下：トウモロコシ）

小麦や牧草は既に生育が進んでいたが，甜菜はまだ小さく，豆類・馬鈴薯・トウモロコシはまだ出芽していないか，出芽したばかりであった。



図 3-4 2 期（2015 年 6 月 19 日・20 日）の各圃場の様子（左上：小麦，右上：小豆，  
左中：甜菜，右中：馬鈴薯，左下：牧草，右下：トウモロコシ）

小麦は完全に生育が進み，甜菜も大幅に生育が進んでいていた。馬鈴薯は一部開花していた。豆類やトウモロコシは草高等に変化があまりなかった。牧草は一番草の刈取直後であった。



図 3-5 3期（2015年7月14日・15日）の各圃場の様子（左上：小麦，右上：小豆，  
左中：甜菜，右中：馬鈴薯，左下：牧草，右下：トウモロコシ）

小麦は登熟期に入り，黄色くなっていた。トウモロコシは大幅に生育が進み，甜菜や馬鈴薯も更に生育が進んでいた。豆類もわずかに生育が進んでいた。牧草は刈取後であった。



図 3-6 4 期（2015 年 8 月 6 日）の各圃場の様子（左上：小麦，右上：小豆，  
左中：甜菜，右中：馬鈴薯，左下：牧草，右下：トウモロコシ）

小麦は収穫に近い状態であった。豆類は開花し始めており，トウモロコシは大幅に生育が進み出穂していた。甜菜も生長がみられた。馬鈴薯は枯凋剤により黄化し始めていた。牧草は二番草の刈取前の状態であった。



図 3-7 5期（2015年8月27日・28日）の各圃場の様子（左上：小麦，右上：小豆，  
左中：甜菜，右中：馬鈴薯，左下：牧草，右下：トウモロコシ）

小麦と牧草は刈取が終了していた。豆類の開花や甜菜の登熟，トウモロコシの出穂が更に進み，生長はピークに達する時期となっていた。馬鈴薯は収穫直前のため，地上部は枯死し始めていた。

### 3.3.2 決定木の数の検討結果

OOB サンプルを使用して、決定木の数について検討を行ったところ、決定木の数とエラー率との関係は図 3-8 のとおりとなった。この結果から、決定木の数がおよそ 50 本以上で OOB サンプルのエラー率が低く安定していることが判明した。分類において、決定木の数が多いことに関しては問題ないとされているため (Breiman, 2001), 決定木の数をそれより大きい  $n=200$  と設定し、残り 80% のデータを分類して精度を評価した。

### 3.3.3 RADARSAT-2 データを使用した分類

まず、RADARSAT-2 データのみを使用して分類を行った。全ての分類結果から得られた OA と  $\kappa$  を図 3-9 に示す。1 データセットのみを使用した分類において、最も高い精度が得られたのは、VZ により算出された要素を使用した場合 (OA=0.918) であった。VZ のみを使用した分類結果の判別効率表を表 3-4 に示す。とくに小麦で PA, UA とともに高く、トウモロコシはいずれも低かった。甜菜は他の圃場に誤分類される傾向がみられたため PA が低く、馬鈴薯は他の圃場を馬鈴薯として誤分類される傾向がみられたため UA が低くなった。

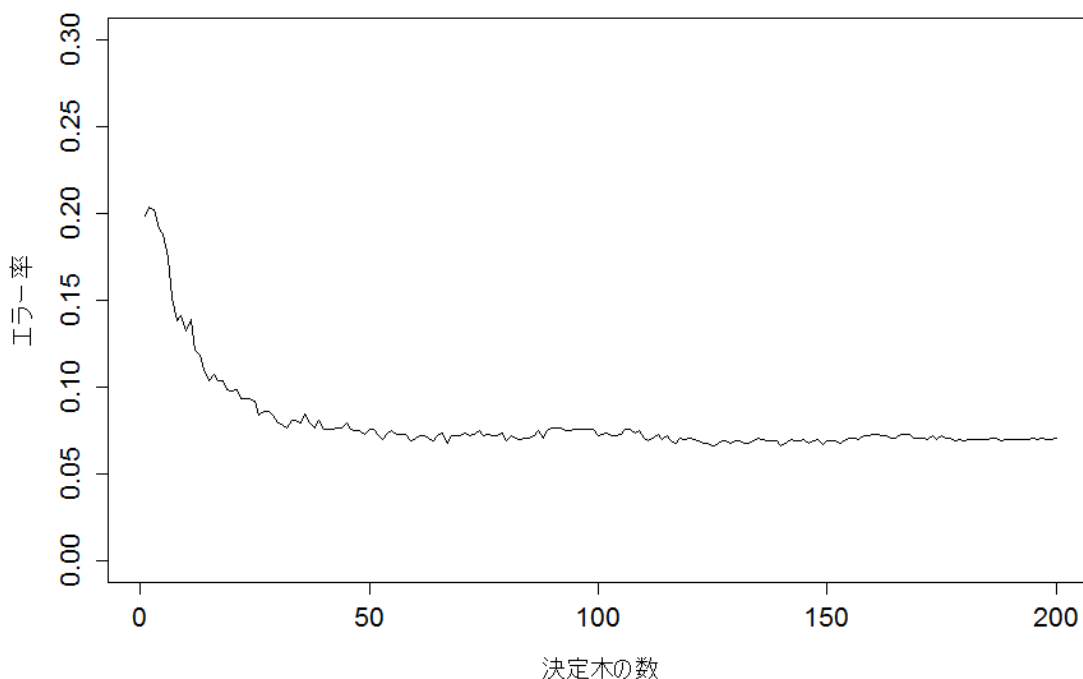


図 3-8 決定木の数とエラー率との関係

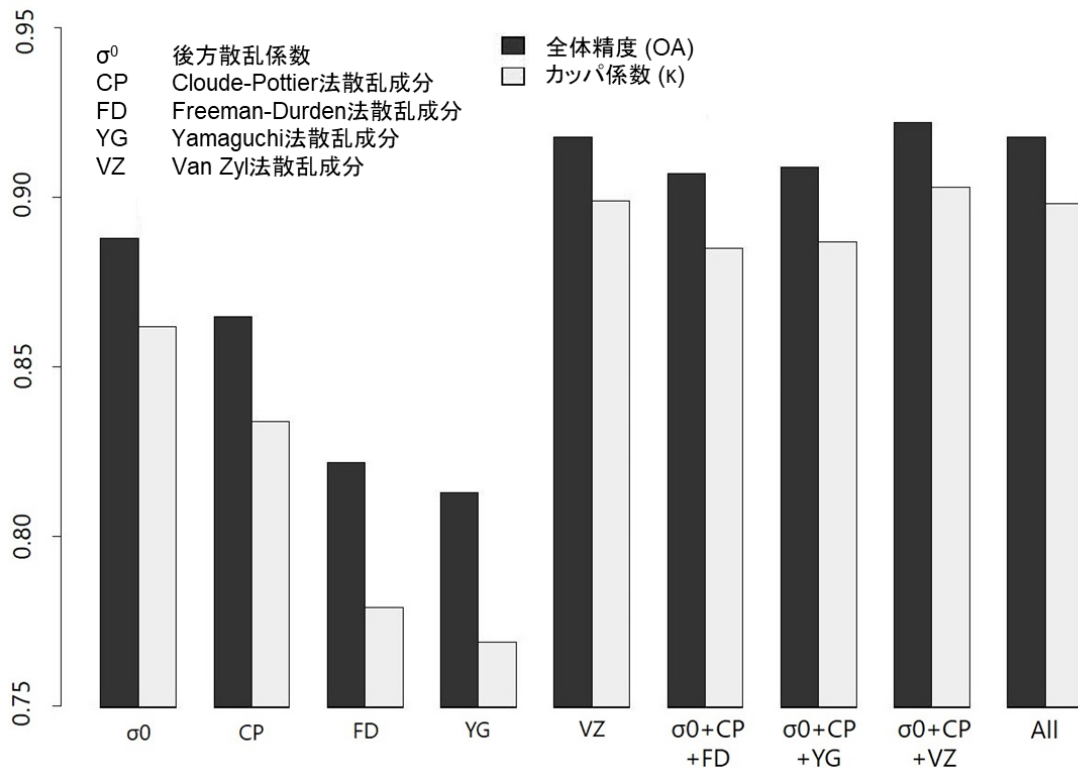


図 3-9 RADARSAT-2 データを使用した各分類結果

表 3-4 RADARSAT-2 の VZ のみを使用した分類結果

参照データ								
	豆類	甜菜	牧草	トウモロコシ	馬鈴薯	小麦	合計	UA
豆類	762	16	5	31	11	2	827	0.921
甜菜	6	396	4	1	5	4	416	0.952
牧草	8	2	479	22	8	19	538	0.890
トウモロコシ	34	12	15	413	20	4	498	0.829
馬鈴薯	17	15	4	30	548	7	621	0.882
小麦	8	3	9	7	3	1,131	1,161	0.974
合計	835	444	516	504	595	1,167	4,061	
PA	0.913	0.892	0.928	0.819	0.921	0.969		
								OA 0.918
								κ 0.899

UA: ユーザーズ精度, PA: プロデューサーズ精度, OA: 全体精度, κ: カッパ係数

1 データセットのみを使用して分類する場合、VZ の散乱成分が作付作物の分類に最も効果的であることが示された。2 章において TerraSAR-X の X バンドデータにより同様の場所、時期数、方法で行われた 1 データセット ( $\sigma^0$  または CD) のみの分類で、CD を用いて 0.841 の OA が得られていた。今回の VZ における結果は、それより 0.07 以上高い精度であった。このことは、C バンド衛星による作付作物分類の有効性を示唆している。一方、VZ 以外のデータセットではいずれも 0.9 を下回り、高い OA を得ることはできなかったが、 $\sigma^0$  と CP では 2 章の X バンドによる結果 (OA = 0.841) より高い精度を得られた。

次に、要素の異なる 3 データセットを組み合わせて分類を行った。その結果、図 3-9 に示すように、OA はそれぞれ 0.9 を超える精度となった。その中でも特に、 $\sigma^0$ 、CP、VZ の組み合わせで分類を実行した場合、最も高い精度 (OA = 0.922) が得られ、 $\kappa$  も 0.9 を上回った。 $\sigma^0$ 、CP、VZ を併用した分類結果の判別効率表を表 3-5 に示す。各作物の PA、UA においては、表 3-4 と同様の傾向がみられた。OA は先述の VZ のデータセットのみを使用した分類結果よりも高い精度であり、算出される要素の異なる 3 データセットを組み合わせることの有用性が示された。類似の研究でも、作付作物の分類において、後方散乱係数、固有値解析法、散乱モデル分解法の 3 グループで算出される 9 要素を併用することの有用性は述べ

表 3-5 RADARSAT-2 データの $\sigma^0$ 、CP、VZ を使用した分類結果

		参照データ							
		豆類	甜菜	牧草	トウモロコシ	馬鈴薯	小麦	合計	UA
分 類 結 果	豆類	762	19	4	35	8	4	832	0.916
	甜菜	7	401	5	2	3	5	423	0.948
	牧草	9	3	484	13	8	14	531	0.911
	トウモロコシ	32	5	15	415	26	3	496	0.837
	馬鈴薯	17	12	4	32	545	5	615	0.886
	小麦	8	4	4	7	5	1,136	1,164	0.976
	合計	835	444	516	504	595	1,167	4,061	
PA		0.913	0.903	0.938	0.823	0.916	0.973		
								OA	0.922
								$\kappa$	0.903

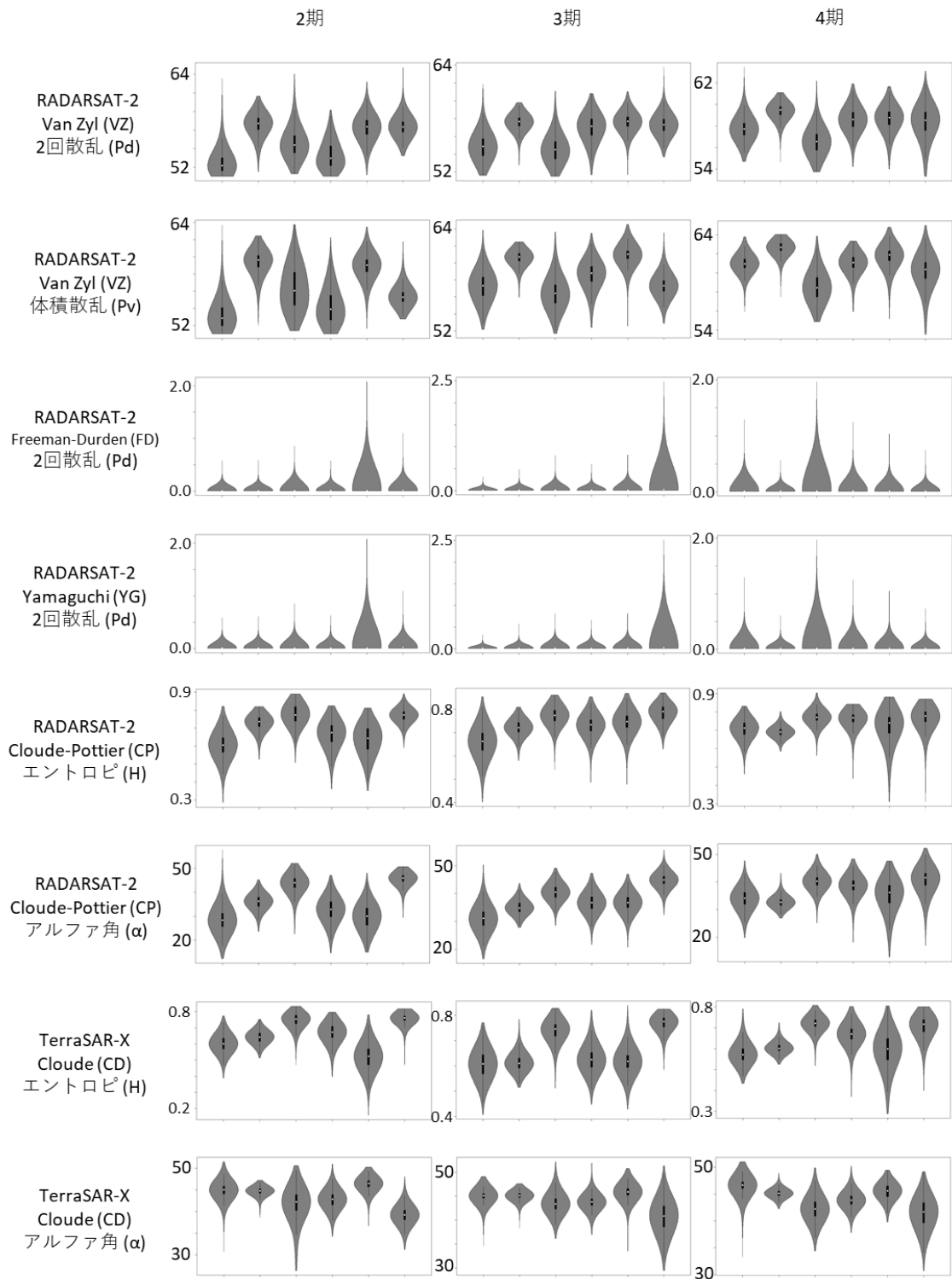
UA: ユーザーズ精度, PA: プロデューサーズ精度, OA: 全体精度,  $\kappa$ : カッパ係数

られており (Xu *et al.*, 2014), 同様な結果が得られたと言える。しかし,  $\sigma^0$ , CP, VZ の 3 データセットを組み合わせた分類と, VZ のみを使用した分類との間には,  $\kappa$  に有意差がなかった。したがって, 散乱モデル分解法である VZ のみを使用することで, 他のデータセットを使用しなくてもある程度高い精度で分類を行えることが示された。

FD や YG の各変数で精度が低下した理由として, 散乱成分の異常値が理由として考えられる。図 3-10 に, 一部の変数における作物別のバイオリンプロットを示す。バイオリンプロットでは, 縦軸で示した変数の値について, その値を持つ圃場の個数を太さで示している。このうち 2 回散乱に着目すると, FD や YG では, VZ と比較して, ほとんどの時期, 作物において圃場ごとの値の分布が 0.0~0.3 に偏っている傾向がみられた。FD や YG の課題として, 植生内において負の値になる場合があるという欠陥が示されており, その課題を解決するために改良された分解手法が VZ である (van Zyl *et al.*, 2011)。本章でも, VZ を使用することでこのような課題が解決され, 分類を行うためのより有用な変数が得られたと考えられる。

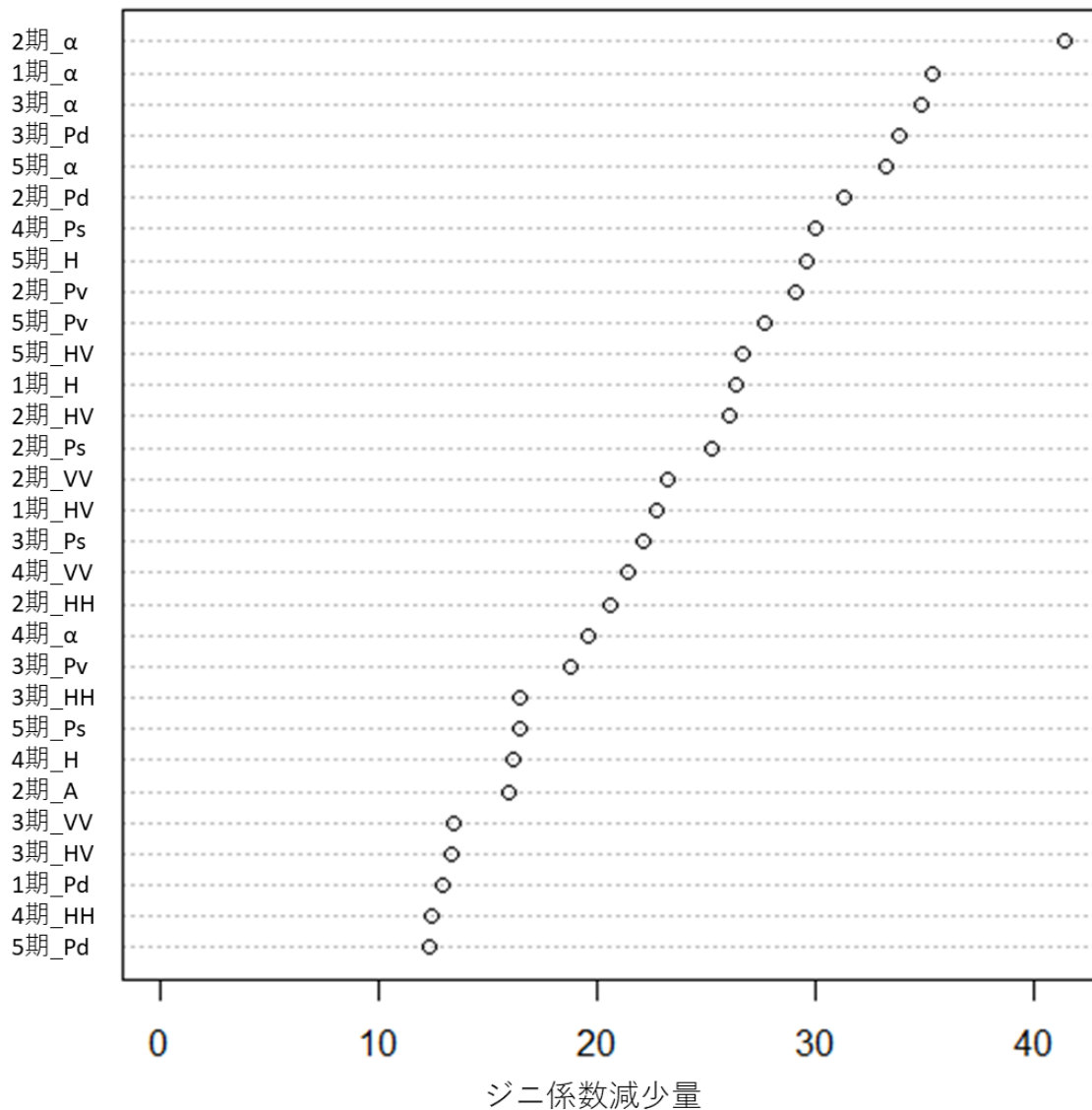
最後に, 5 種類すべてのデータセットを組み合わせて使用した分類を行った (図 3-9 右端)。その結果, OA は 0.918 となり,  $\sigma^0$ , CP, VZ の 3 データセットのみを組み合わせた場合と有意な差はみられなかっただけでなく, むしろ OA や  $\kappa$  は劣っていた。この要因として, 類似した要素を算出する散乱成分分解 (FD, VZ, YG) 間に多重共線性があることが考えられる。併用しても精度の向上が期待できないだけでなく, 今回の結果のように精度の低下を招く可能性もあるため, 併用する必要性は低いといえる。

最も精度の高かったデータセットの組み合わせ ( $\sigma^0$ , CP, VZ) での分類について, ジニ係数減少量より変数の重要度の評価を行った (図 3-11)。その結果, 1~3, 5 期のアルファ角や, 2, 3 期の 2 回散乱で特に高い重要度を示したほか, 一部の体積散乱やエントロピでも高い重要度を示した。2, 3 期にあたる 6~7 月にかけては, 作物ごとに生育の段階や作物体の大きさが異なる。図 3-10 の上から 6 番目に示した, 2~4 期のアルファ角における作物別のバイオリンプロットからもみられるように, アルファ角も作物間で差異が大きく, 分類における精度の向上に寄与したと考えられる。また, 同図の上から 1 番目に示した 2~4 期の 2 回散乱における作物別のバイオリンプロットより, 2 回散乱は小麦のような垂直構造を持つ作物で他の作物と差異が大きくなるため, 6~7 月の小麦が出穂・開花し始めている 2~3 期で特に高い重要度を示したと考えられる。また, 作物体が大きくなり構造が複雑になると



各プロットの横軸は左から豆類, 甜菜, 牧草, トウモロコシ, 馬鈴薯, 小麦

図 3-10 作物別各変数のバイオリンプロット (一部)



HH: HH偏波後方散乱係数    HV: HV偏波後方散乱係数    VV: VV偏波後方散乱係数  
 Pd: 2回散乱    Pv: 体積散乱    Ps: 表面散乱  
 H: エントロピ    A: アニソトロピ    α: アルファ角

図 3-11 RADARSAT-2 データの $\sigma^0$ , CP, VZ を使用した分類におけるジニ係数減少量

体積散乱が発生しやすくなるため、8月のような作物の生育ごとに差異が大きい時期の重要度が高くなったと考えられる。

以上のように、VZ や CP の散乱成分は特に作付作物を把握するための重要な変数となりうると考えられる。既往研究においても、後方散乱係数、FD の散乱モデル分解法および固

有値による散乱成分の計 9 変数を使用することで、作付作物分類を高精度に行えることが示されている (Jiao *et al.*, 2014)。一方、本章では散乱モデル分解法として VZ を用いることにより、更なる精度の向上が示された。かつ、体積散乱や 2 回散乱のような作物の構造によって値に変化が大きくなる変数は特に重要度が高くなったことから、作付作物分類における VZ の有用性が示された。

### 3.3.4 RADARSAT-2 データと TerraSAR-X データとを組み合わせた分類

次に、RADARSAT-2 データと TerraSAR-X データとを組み合わせて分類を行った。全ての分類結果から算出された OA と  $\kappa$  を図 3-12 に示す。RADARSAT-2 の 3 データセット ( $\sigma^0$ , CP, VZ) を RS2 とし、TerraSAR-X の両方のデータセット ( $\sigma^0$ , CD) を TSX としている。

RADARSAT-2 データに TerraSAR-X のデータセットのうちいずれかを併用した分類結果 ( $\sigma^0$ : OA=0.928,  $\kappa$ =0.911, CD: OA=0.929,  $\kappa$ =0.912) は、どちらも RADARSAT-2 のみでの分類結果 (OA=0.922,  $\kappa$ =0.903) より高い精度となった。ただし図 3-12 に示した各分類結果間における Z 検定より、これらの  $\kappa$  の間には 5% の有意水準で有意な差がなく、RADARSAT-2 データに TerraSAR-X データからの 1 データセットを併用して分類を行う必要性は低い。TerraSAR-X の 1 データセットのみを併用した作物別の分類による各精度指標を表 3-6 に示す。作物別に見た場合、いずれも小麦で高い PA, UA を示し、トウモロコシで PA, UA が低くなるという結果がみられた。この理由として、小麦は他の作物が出芽して間もない 5 月下旬~6 月には秋播き小麦で出穂が近い時期となっているほか、春播き小麦でも作物体が大きくなっている点が挙げられる (図 2-2)。そのため、各波長におけるマイクロ波の散乱も他の作物と大幅な差異があったと考えられる。一方、トウモロコシは同作物の圃場間によっても各変数の値が大きく異なることがあった。これは、デントコーンとスイートコーンで生育状態に差異がある点や、同じスイートコーン圃場でも収穫期をずらすために生育状態に差異がある点などが考えられる。したがって、圃場によって各変数のばらつきが大きく、同じ時期に生育が進んでいる豆類や馬鈴薯との間で誤分類が多く生じたことが示唆される (図 2-2)。また、 $\sigma^0$  を併用した場合と CD を併用した場合で比較すると、両者の分類結果に有意な差はなかった。

TerraSAR-X から算出された 2 種類すべてのデータセットを RADARSAT-2 と併用したと

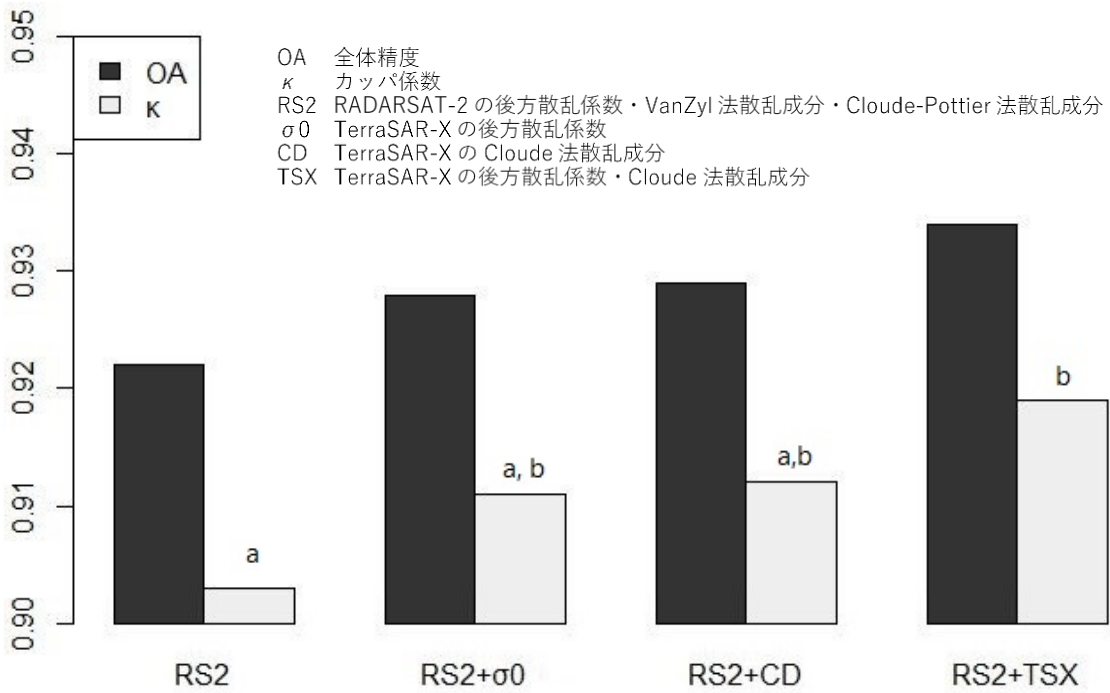


図 3-12 RADARSAT-2 データと TerraSAR-X データを併用した各分類結果

(カラム上のアルファベット a, b は Z 検定の結果を示しており、  
同じ文字の κ 間には有意な差がないことを示す)

表 3-6 RADARSAT-2 データの $\sigma^0$ , CP, VZ および TerraSAR-X の 1 データセットを  
併用した分類における各種精度指標

	RS2+ $\sigma^0$		RS2+CD	
	PA	UA	PA	UA
豆類	0.921	0.925	0.919	0.912
甜菜	0.923	0.960	0.928	0.948
牧草	0.940	0.915	0.944	0.912
トウモロコシ	0.817	0.862	0.843	0.885
馬鈴薯	0.919	0.881	0.926	0.903
小麦	0.982	0.976	0.975	0.973
OA		0.928		0.929
κ		0.911		0.912

RS2: RADARSAT-2 の後方散乱係数, van Zyl 法散乱成分, Cloude-Pottier 法散乱成分の計 9 変数  
 $\sigma^0$ : TerraSAR-X の後方散乱係数, CD: TerraSAR-X の Cloude 法散乱成分  
 UA: ユーザーズ精度, PA: プロデューサーズ精度, OA: 全体精度, κ: カッパ係数

ころ、片方のデータセットのみを使用した場合よりさらに高い精度が得られた (OA=0.934)。この結果は RADARSAT-2 のみで分類を行った場合と比較して、 $\kappa$  には 5%の有意水準で有意な差があった。したがって、RADARSAT-2 の3データセットに TerraSAR-X データの2データセットを併用した分類は、いずれかの衛星のみを使用した場合と比較して、作物や地表面での透過や応答が異なるために、併用する意義があったといえる。この結果の判別効率表を表 3-7 に示す。表 3-7 から、各作物で PA や UA の向上がみられた。表 3-6 に示したとおり、TerraSAR-X の1データセットのみを併用した場合、 $\sigma^0$ では馬鈴薯やトウモロコシの PA, UA が低い一方で、CD では豆類や甜菜の UA が低いなど、それぞれの結果で PA, UA の低い作物が異なっていた。したがって、特徴の異なる双方のデータセットを RADARSAT-2 に併用することで情報を補完し合い、より作物を判別しやすい決定木が作成され、精度が向上したと考えられる。

最も精度の高かったデータセットの組み合わせ (RADARSAT-2 の $\sigma^0$ , CP, VZ および TerraSAR-X の $\sigma^0$ , CD) での分類について、ジニ係数減少量から重要度の評価を行った (図 3-13)。その結果、2, 3期の TerraSAR-X のアルファ角と3期の TerraSAR-X のエントロピで高い重要度を示した。図 3-10 の下2つに示した TerraSAR-X の2~4期におけるアルファ角およびエントロピにおける作物別のバイオリンプロットによると、TerraSAR-X でも

表 3-7 RADARSAT-2 の $\sigma^0$ , CP, VZ および TerraSAR-X の $\sigma^0$ , CD を併用した分類結果

参照データ								
	豆類	甜菜	牧草	トウモロコシ	馬鈴薯	小麦	合計	UA
豆類	778	13	5	30	18	4	848	0.917
甜菜	8	408	2	3	1	3	425	0.960
分 牧草	9	2	491	15	9	4	530	0.926
類 トウモロコシ	13	11	9	421	12	3	469	0.898
結 馬鈴薯	16	7	5	28	551	7	614	0.897
果 小麦	11	3	4	7	4	1,146	1,175	0.975
合計	835	444	516	504	595	1,167	4,061	
PA	0.932	0.919	0.952	0.835	0.926	0.982		
								OA 0.934
								$\kappa$ 0.919

UA: ユーザーズ精度, PA: プロデューサーズ精度, OA: 全体精度,  $\kappa$ : カッパ係数

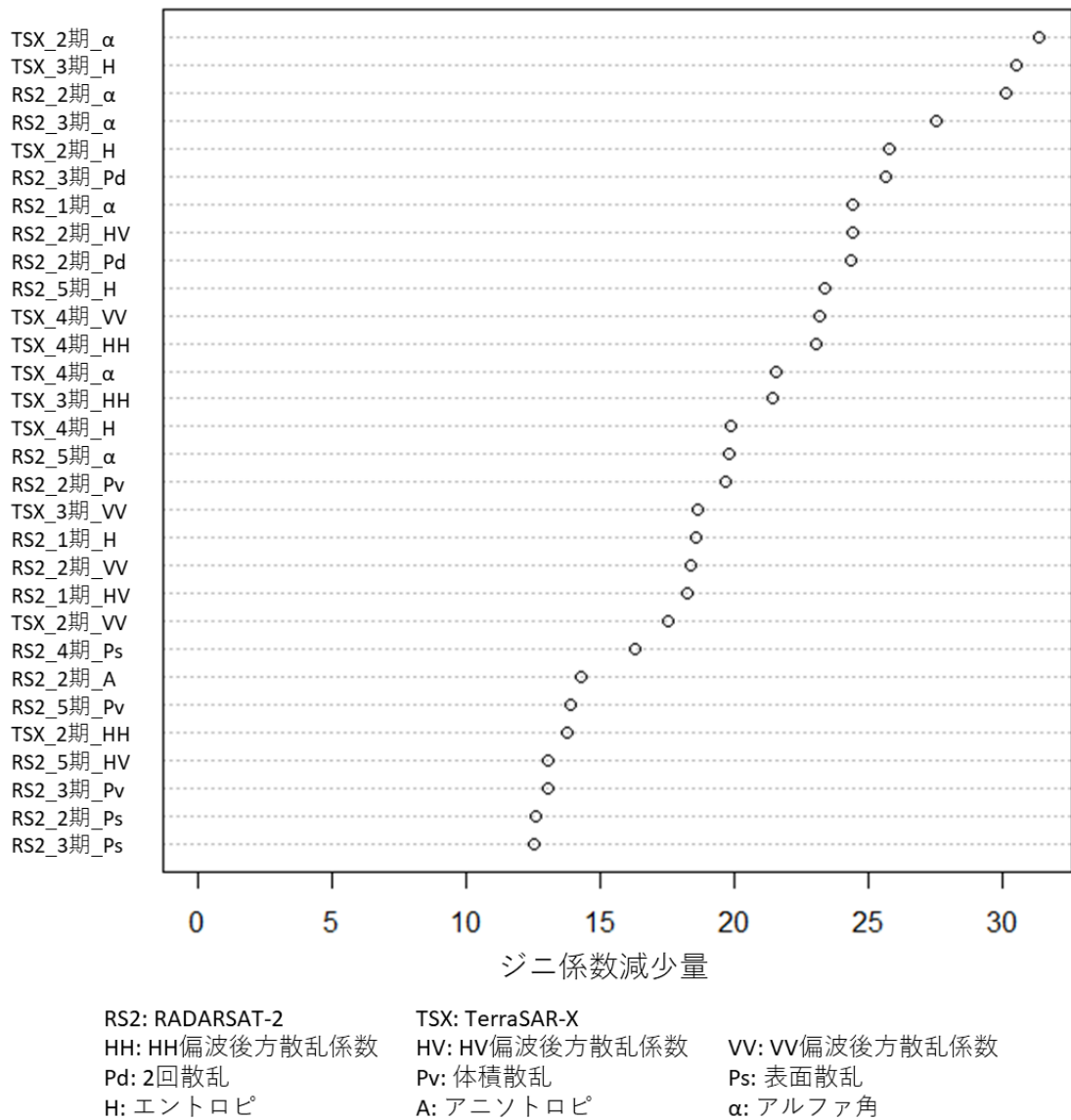


図 3-13 RADARSAT-2 データの $\sigma^0$ , CP, VZ および TerraSAR-X データの $\sigma^0$ , CD を併用した分類におけるジニ係数減少量

RADARSAT-2 と同様に、作物ごとに生育段階や作物体の大きさが異なる 6~7 月 (2, 3 期) において、これらの値がそれぞれ異なっていた。この時期は TerraSAR-X においても作物ごとに散乱の形態が異なるために、エントロピやアルファ角の値も異なりやすく、分類に効果があったと考えられる。また、3.3.1 の結果と同様に、RADARSAT-2 の 1~3 期のアルファ角や、2, 3 期の 2 回散乱も高い重要度を示した。TerraSAR-X のデータを加えても

RADARSAT-2 のこれらの変数で高い重要度を示したことや、図 3-10 で示したように 3~4 期におけるエントロピの値が同じ作物でも波長間で傾向が異なっていたことから、TerraSAR-X と RADARSAT-2 では作物体での透過の度合いや散乱形態が異なるために、それぞれの変数が有効となることが示唆された。したがって、畑地における作付作物分類において、両者を併用することの意義が示された。また、TerraSAR-X では、4 期の変数でも高い重要度を示した。これは、TerraSAR-X の方が短波長で作物体を透過しにくく、多くの作物が繁茂している時期である 8 月は土壌の影響を受けにくいために、各作物の特徴を捉えやすかったことが考えられる。

### 3.3.5 使用する時期数の検討

3.3.3 で述べた RADARSAT-2 データのみを使用した分類において、同じ組み合わせのデータセットで、使用する時期数や時期の組み合わせを変えて再度分類を行い、OA や  $\kappa$  の差異を評価した。3.3.3 では、 $\sigma^0$ 、CD、VZ を使用した分類が最も高い精度であったため、ここでもこの 3 データセットを使用した分類で時期数の検討を行った。表 3-8 に、各時期数で使ったデータの観測時期を示す。各時期数において用いた時期は、最も高い精度を示した時期の組み合わせである。表 3-8 に示したとおり、少ない時期数で選択された時期は、時期数を増やしても必ず残り、入れ替わることはなかった。表 3-9 に、各時期数における精度の各指標を示す。

1 時期のみを使用して分類を実行したところ、最高の精度となったのは 3 期 (7 月 13 日)

表 3-8 RADARSAT-2 の  $\sigma^0$ , CP, VZ を使用した時期数別の分類における

使用した観測時期の組み合わせ					
時期数	使用した時期				
	1期 5月26日	2期 6月19日	3期 7月13日	4期 8月06日	5期 8月30日
1時期			○		
2時期			○		○
3時期		○	○		○
4時期	○	○	○		○
5時期	○	○	○	○	○

表 3-9 RADARSAT-2 の $\sigma^0$ , CP, VZ を使用した時期数別の分類における各種精度指標

	作物	時期数				
		1時期	2時期	3時期	4時期	5時期
PA	豆類	0.859	0.887	0.904	0.908	0.913
	甜菜	0.493	0.845	0.863	0.899	0.903
	牧草	0.806	0.903	0.921	0.946	0.938
	トウモロコシ	0.619	0.760	0.794	0.786	0.823
	馬鈴薯	0.714	0.852	0.904	0.916	0.916
	小麦	0.927	0.950	0.973	0.973	0.973
UA	豆類	0.821	0.878	0.898	0.906	0.916
	甜菜	0.590	0.874	0.927	0.921	0.948
	牧草	0.800	0.903	0.926	0.900	0.911
	トウモロコシ	0.629	0.723	0.795	0.828	0.837
	馬鈴薯	0.664	0.865	0.875	0.898	0.886
	小麦	0.932	0.959	0.966	0.975	0.976
OA		0.781	0.882	0.908	0.916	0.922
$\kappa$		0.729	0.854	0.886	0.897	0.903

UA: ユーザーズ精度, PA: プロデューサーズ精度, OA: 全体精度,  $\kappa$ : カッパ係数

のみを使用した場合 (OA: 0.781,  $\kappa$ : 0.729) であった (表 3-8, 3-9)。この時期は、小麦が登熟し始める一方で牧草は2番草の収穫をしており、豆類や甜菜、トウモロコシは生長段階にある時期である。このように、圃場ごとに作物体の大きさが多様性に富んでいる時期であるため、1時期のみを使用した分類の中でも、この時期は他の時期より高い精度での分類が可能であったと考えられる。しかし、1時期のみを使用した分類では、どの時期であっても0.80のOAを超えることはなかった。

2時期を使用して分類を実行したところ、最高の精度となったのは3期に5期を加えた場合 (OA: 0.882,  $\kappa$ : 0.854) であった (表 3-8, 3-9)。この結果から、2時期を使用することで精度の大幅な向上に寄与するといえる。表 3-9 に示したとおり、 $\kappa$  は 0.8 を上回っており、OA も 1 時期のみを使用して分類を実行した場合と比較して 0.1 以上の上昇がみられた。小麦や牧草が収穫後、豆類が開花後、馬鈴薯が収穫直前、甜菜が登熟後、トウモロコシが出穂後と、圃場の状態が作物によって異なる 8 月下旬のデータを 7 月中旬のデータと組み合わせるこ

とによって、高い精度が得られたと考えられる (図 2-2)。作物別では、甜菜、馬鈴薯、トウモロコシ、牧草で精度が 1 時期を用いた分類と比較して 0.1 以上上昇しており、中でも甜菜は PA, UA とも大幅に上昇した。

3 時期を使用して分類を実行したところ、最高の精度となったのは 3,5 期に 2 期を加えた場合 (OA: 0.908,  $\kappa$ : 0.886) であり (表 3-8, 3-9), 3 時期を使用することで初めて OA が 0.9 を上回った。2 時期を使用して分類を実行した場合と比較したところ、両者の  $\kappa$  に有意な差が認められた。すなわち、3 時期を使用して分類することで、より高精度での分類が可能となるといえる。作物別では、馬鈴薯の PA やトウモロコシの UA が特に上昇したが、大幅に精度上昇がみられた作物はなかった。

4 時期を使用して分類を実行したところ、最高の精度となったのは 2,3,5 期に 1 期を加えた場合 (OA: 0.916,  $\kappa$ : 0.897) であった (表 3-8, 3-9)。4 時期を使用した分類は、3.3.3 で示した 5 時期すべてを使用した分類 ( $\kappa$ : 0.903) と比較して、 $\kappa$  に有意な差は認められなかった。すなわち、4 時期のみを使用した分類は、5 時期すべてを使用した分類と同じ水準の精度を得られるといえる。

RADARSAT-2 データと TerraSAR-X データを併用した分類でも、RADARSAT-2 のみを使用した分類と同様に、使用する時期数や時期の組み合わせを変えて再度分類を行い、OA や  $\kappa$  の差異を評価した。3.3.4 で述べたように、RADARSAT-2 の  $\sigma^0$ , CP, VZ と TerraSAR-X の  $\sigma^0$ , CD を使用した分類が最も高い精度であったため、ここでもこのデータセットを併用し

表 3-10 RADARSAT-2 の  $\sigma^0$ , CP, VZ および TerraSAR-X の  $\sigma^0$ , CD を併用した時期数別の分類における使用した観測時期の組み合わせ

時期数	使用した時期				
	1期	2期	3期	4期	5期
	RADARSAT-2 5月26日	6月19日	7月13日	8月06日	8月30日
TerraSAR-X	5月22日	6月24日	7月16日	8月07日	8月29日
1時期		○			
2時期		○		○	
3時期		○	○	○	
4時期		○	○	○	○
5時期	○	○	○	○	○

た分類で時期数の検討を行った。表 3-10 に、各時期数で使用したデータの観測時期を示す。こちら、各時期数において用いた値は、最も精度の高い時期を組み合わせたものである。表 3-10 に示したとおり、少ない時期数で選択された時期は、時期数を増やしても必ず残り、入れ替わることはなかった。表 3-11 に、各時期数における精度の各指標を示す。

1 時期のみを使用して分類を実行したところ、最高の精度となったのは 2 期のみを使用した場合 (OA: 0.870,  $\kappa$ : 0.840) であった (表 3-10, 3-11)。ほぼすべての作物が登熟、収穫前の時期で最高の精度が得られたのは、RADARSAT-2 のみを使用した 1 時期のみでの分類結果でもほぼ同様の傾向であった (表 3-8, 3-9)。しかし、OA そのものは RADARSAT-2 データのみを使用した場合と比較して、0.1 近く上昇していた (表 3-9, 3-11)。そのため、1 時期のみのデータであっても、2 種類の波長を使用することで、精度を向上させることができることが示された。この理由として、作物をある程度透過する RADARSAT-2 だけでなく、作

表 3-11 RADARSAT-2 の $\sigma^0$ , CP, VZ および TerraSAR-X の $\sigma^0$ , CD を併用した  
時期数別の分類における各種精度指標

	作物	時期数				
		1時期	2時期	3時期	4時期	5時期
PA	豆類	0.871	0.923	0.923	0.927	0.932
	甜菜	0.831	0.903	0.896	0.926	0.919
	牧草	0.893	0.917	0.936	0.938	0.952
	トウモロコシ	0.635	0.819	0.835	0.827	0.835
	馬鈴薯	0.918	0.914	0.924	0.929	0.926
	小麦	0.952	0.973	0.974	0.980	0.981
	UA	豆類	0.822	0.915	0.916	0.921
甜菜		0.937	0.941	0.954	0.967	0.960
牧草		0.829	0.917	0.917	0.934	0.925
トウモロコシ		0.726	0.857	0.863	0.871	0.896
馬鈴薯		0.857	0.880	0.883	0.886	0.899
小麦		0.967	0.965	0.977	0.974	0.975
OA		0.870	0.920	0.926	0.932	0.934
$\kappa$	0.840	0.901	0.908	0.915	0.919	

UA: ユーザーズ精度, PA: プロデューサーズ精度, OA: 全体精度,  $\kappa$ : カッパ係数

物の表面で主に散乱する TerraSAR-X の変数を決定木に使用することで (石塚・大内, 2017), 作物体の判別がより詳細にできるようになったためであると考えられる。

2 時期を使用して分類を実行したところ, 最高の精度となったのは 2 期に 4 期を加えた場合 (OA: 0.920,  $\kappa$ : 0.910) であった (表 3-10, 3-11)。この結果から, RADARSAT-2 データのみを使用した分類と同様に, 2 時期を使用することで大幅な精度の向上がみられただけでなく, 2 時期のみの使用で OA が 0.9 を上回った。くわえて, 2 時期を使用した分類では, 1 時期のみを使用した分類と比較したところ, 両者の  $\kappa$  に有意な差がみられたことから, より良好な結果を得られたといえる。また, この精度は, RADARSAT-2 データのみを使用した場合の分類では 4 時期分にあたり, 半分の時期数で同等の精度の分類が可能となった。以上の結果から, RADARSAT-2 データと TerraSAR-X データとを組み合わせる分類を行う場合には, 6 月下旬と 8 月上旬の衛星データを取得することで, 各作物の詳細な分類が可能となることが示された。

3 時期を使用して分類を実行したところ, 最高の精度となったのは 2, 4 期に 3 期を加えた場合 (OA: 0.926,  $\kappa$ : 0.908) であった (表 3-10, 3-11)。同様に 4 時期を使用して分類を実行したところ, 最高の精度となったのは 2, 3, 4 期に 5 期を加えた場合 (OA: 0.932,  $\kappa$ : 0.915) であり, 4 時期を使用することで 0.93 の OA を達成することが可能となった。4 時期を使用した分類は, 3.2 で示した 5 時期すべてを使用した分類 (OA: 0.934,  $\kappa$ : 0.919) と比較して, 両者の  $\kappa$  に有意な差は認められなかった。したがって, RADARSAT-2 データと TerraSAR-X データとを併用して分類を行う場合, 5 時期のデータを使用しなくても, 3 時期や 4 時期のデータ数でもほぼ同じ水準での分類が可能であることが示された。4 時期のデータと 5 時期のデータを比較したところ, 甜菜の PA や UA, 牧草の UA などは, 5 時期での分類で 4 時期よりも低下する傾向がみられた (表 3-11)。これは, 1 期ではこれらの作物体がまだ小さいことから, 他の作物と判別するための変数として不適切であることが考えられる。これらの変数を分類に含めることにより, RF における決定木の過適合が発生したと考えられる。すなわち, 分類に使用する変数の数が多ければよいというわけではなく, 作物体を判別するための重要な時期のデータのみを予め選択する重要性も示唆される。

同じ時期数で比較した場合, RADARSAT-2 のみによる分類と比較して, RADARSAT-2 と TerraSAR-X を併用した分類の方が, 選択される時期が早まる傾向がみられた (表 3-8, 3-10)。また, RADARSAT-2 の  $\sigma^0$ , CP, VZ および TerraSAR-X の  $\sigma^0$ , CD を用いた 2, 3, 4 期での

分類について、ジニ係数減少量により重要度の評価を行った。その結果、RADARSAT-2からは2, 3期のアルファ角や3期の2回散乱で高い重要度を示した。TerraSAR-Xからは2期のアルファ角やエントロピ、4期のHH, VVで高い重要度を示した。これらの理由として、小麦以外の作物で作物体の生育が進んでいる6月下旬, 7月中旬(2, 3期)で、作物ごとの散乱形態に波長間で差が大きく、分類の精度向上に寄与しやすかったことが考えられる。本地域において、2期は小麦が出穂後かつ、豆類が出芽後、牧草が1番草収穫後であり、3期は小麦が登熟前後、馬鈴薯が開花後である。このように、各作物が最も形態的に多様性に富む時期である。くわえて、3.2と同様に4期の $\sigma^0$ が高い重要度を示した点からも、TerraSAR-Xでは作物体が繁茂した圃場でも各作物の特徴を捉えることができると考えられる。本地域では、4期は小麦が収穫前後、豆類および馬鈴薯が開花後、トウモロコシが出穂後である。すなわち、両者の衛星データを併用することで、高精度な分類マップの作成が可能であるといえる。

最後に、RADARSAT-2データとTerraSAR-Xデータを3時期(2, 3, 4期)使用した分類結果のマップの一部を図3-14に示す。誤分類された圃場を黒で示し、誤分類が生じた圃場の特徴について考察を行った。誤分類が生じた理由として、次の3点が考えられる。

1点目は、ピクセル数が少ない圃場の存在である。誤分類された圃場の中には、狭小な圃場がいくつか散見された。これらの圃場は、ピクセル数が少ないために、SARのノイズの影響を受けた場合に圃場の平均値が外れ値となりやすい。その影響で誤分類が生じやすかった可能性が考えられる。作成した圃場ポリゴンの面積と、誤分類圃場数およびOAの関係を表3-12に示す。70a以上では概ね0.90以上のOAが得られ、150a以上であれば概ね0.98以上のOAが得られた。一方で10a以上では0.70を超えるが、10a未満の圃場ではOAが0.50を下回り、103圃場中53圃場が誤分類される結果となった。すなわち、ポリゴンが10a以下の圃場では、ピクセルが少なすぎるために作付作物分類は困難であるといえる。

2点目は、同じ作物でも圃場間で生育段階のずれがあった可能性がある点である。天候不順や収穫時期をずらすなどのために、同じ作物であればすべての圃場で同一の日時に播種や収穫等を行うわけではない。とくに対象地域においては、スイートコーンや豆類でそのような圃場が多い。また、収穫中の時期に衛星データが観測された場合、特に変数の差異は圃場ごとで大きくなる。このような圃場ごとの生育や植物体の量の違いで誤分類を引き起こ

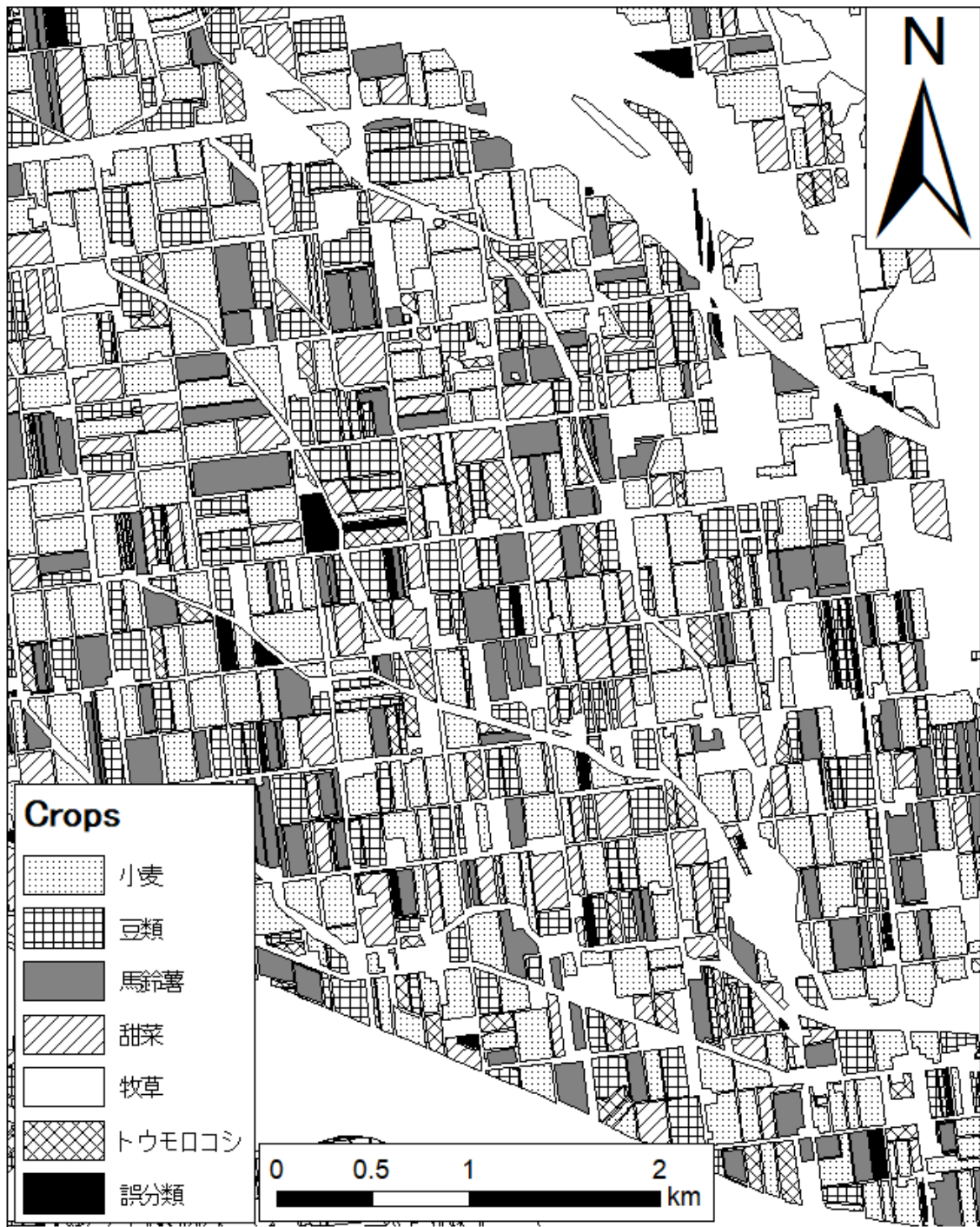


図 3-14 RADARSAT-2 データの $\sigma^0$ , CP, VZ および TerraSAR-X データの $\sigma^0$ , CD を  
3 時期ずつ併用した分類マップ (一部)

表 3-12 RADARSAT-2 の $\sigma^0$ , CP, VZ および TerraSAR-X の $\sigma^0$ , CD を 3 時期ずつ

併用した分類における圃場ポリゴンの面積別全体精度

ポリゴン面積 (a)	テストデータ圃場数	誤分類圃場数	全体精度 (OA)
0 - 10	103	53	0.485
10 - 20	91	26	0.714
20 - 30	99	26	0.737
30 - 40	111	17	0.847
40 - 50	93	20	0.785
50 - 60	115	13	0.887
60 - 70	125	16	0.872
70 - 80	166	12	0.928
80 - 90	143	8	0.944
90 - 100	129	15	0.884
100 - 150	637	29	0.954
150 - 200	620	12	0.981
200 - 250	462	10	0.978
250 - 300	315	6	0.981
300 - 350	223	5	0.978
350 - 400	173	2	0.988
400 - 450	197	4	0.980
450 - 500	81	2	0.975
500 -	178	2	0.989

した可能性も考えられる。以上 2 点は、先述の各時期で生育段階に近い 3 作物（豆類、馬鈴薯、トウモロコシ）の精度低下に影響を及ぼした要因となることが考えられる。

3 点目は、特定の圃場の状態が部分的に他の圃場と異なっていた可能性がある点である。誤分類された圃場の一部には広大な圃場も散見されたが、これらの多くは牧草やトウモロコシの圃場であった。対象地域の牧草地の中には、樹木が植えられている圃場や、収穫後の牧草をロールにして圃場内に放置してある圃場がみられた。これらがある圃場で、特に誤分類が生じた可能性が考えられる。トウモロコシについては、特にデントコーン圃場で比較的広大な圃場が多い。デントコーンや牧草などの飼料作物は、その他の作物と比較して生育ムラの大きい圃場が対象地域には多く存在した。このような生育ムラも圃場内や各圃場での散乱の差異に大きく影響を及ぼし、同時に誤分類を引き起こす原因となったと考えられる。

また、小麦に関しても、風などの影響で登熟期以降に倒伏のみられる圃場がある。倒伏することにより散乱の形態が変化し、誤分類を生じる要因となったと考えられる。

### 3.4 まとめ

本章では、Cバンド SAR である RADARSAT-2 の 4 偏波 SAR データおよび Xバンド SAR である TerraSAR-X の 2 偏波データを各 5 時期ずつ利用し、機械学習アルゴリズムの Random Forests による作付作物の分類を行った。

まず、RADARSAT-2 データのみを使用し、作付作物の分類を行った。その結果、5 時期のデータを使用することにより、圃場の作付作物を 0.9 以上の高い全体精度 (OA) で分類できることが示された。この結果は、2 章で示した TerraSAR-X データのみを使用した分類結果よりも高い精度であり、RADARSAT-2 データを使用して作付作物の分類を行うことの有効性が示された。また、適用するデータセットによる精度の比較も行った。1 データセットでの分類では、van Zyl 法 (VZ) の散乱成分を使用することで最も高い OA が得られた。3 データセットでの分類では、この VZ に後方散乱係数 ( $\sigma^0$ ) や固有値解析法の Cloude-Pottier 法 (CP) を加えて分類を行うことにより、5 データセットすべてを用いた分類よりも精度の高い 0.922 の OA が得られた。以上より、5 時期の RADARSAT-2 データを用いて作付作物の分類を行うにあたっては、 $\sigma^0$ 、CP、VZ の 3 データセットを用いることが最も有効であるといえる。

次に、RADARSAT-2 データおよび TerraSAR-X データを併用し、作付作物の分類を行った。RADARSAT-2 の  $\sigma^0$ 、CP、VZ に、TerraSAR-X の  $\sigma^0$  または固有値解析法の Cloude 法 (CD) のいずれかを加えて精度を比較した。その結果、RADARSAT-2 の 4 偏波 SAR データのみを用いた場合を上回る 0.929 の OA が得られた。しかしこの結果は、RADARSAT-2 データのみを使用した場合と  $\kappa$  に有意差はなかった。そこで、RADARSAT-2 の  $\sigma^0$ 、CP、VZ に、TerraSAR-X の  $\sigma^0$  と CP の両方を加えて精度を比較した。その結果、OA は 0.934 となり、RADARSAT-2 のみを使用した場合と比較して有意な差のある  $\kappa$  が得られた。以上より、5 時期の RADARSAT-2 データおよび TerraSAR-X データを用いて作付作物の分類を行うにあたっては、前者からは  $\sigma^0$ 、CP、VZ の 3 データセットを、後者からは  $\sigma^0$ 、CD の 2 データセットを

それぞれ用いることが最も有効であるといえる。この手法では、いずれか一方の衛星データを使用した場合よりも高い OA で分類可能であったことから、波長が違う衛星データを組み合わせたことにより補完され精度が上がったと考えられる。

最後に、RADARSAT-2 データを使用した分類と、RADARSAT-2 データおよび TerraSAR-X データを併用した分類のそれぞれの結果について、5 時期すべての衛星データを使用しなくても、同様の高精度な分類が可能であるかどうかを検討した。RADARSAT-2 データの $\sigma^0$ , CP, VZ を使用した分類では、3 時期以上を使用することにより、0.9 を超える OA が得られた。さらに、4 時期のデータを使用することにより、5 時期すべてのデータを使用した場合と有意な差のない  $\kappa$  が得られた。RADARSAT-2 データの $\sigma^0$ , CP, VZ と、TerraSAR-X データの $\sigma^0$ , CD を併用した分類では、2 時期のみの使用により 0.9 を超える OA が得られた。さらに、3 時期のデータであれば、5 時期すべてを使用した場合と比較して有意な差のない  $\kappa$  を得ることができた。以上のように、取得する時期数を削減して作付作物の分類を行うにあたっては、RADARSAT-2 のみであれば 4 時期のデータを使用することが最も有効であるといえる。また、RADARSAT-2 と TerraSAR-X の併用であれば 3 時期のデータを使用することが最も有効であるといえ、RADARSAT-2 のみの分類よりも精度が高い結果を得られる。なお、ここで使用するデータの 3 時期の目安として、次に示すものを使用することで特に高い精度となる。1) 小麦出穂後、豆類出芽後、馬鈴薯開花前後、牧草 1 番草収穫後。本地域では 6 月下旬。2) 小麦登熟前後、馬鈴薯開花後、牧草 2 番草収穫後。本地域では 7 月中旬。3) 小麦収穫前後、豆類開花後、トウモロコシ出穂後。本地域では 8 月上旬。また、本章で使用した圃場ポリゴンは、ミクセルの影響を考慮して 10m のバッファを取り、その領域を除いた。その際、除いた後のポリゴンの大きさが 10a 以下の圃場では誤分類が生じやすいため、狭小な圃場での分類は困難であるといえ、注意が必要である。逆に、1.5ha 以上の圃場であれば、95%以上の精度で分類を行うことができる。



## 4 SAR を利用した高精度な作付作物分類のための機械学習アルゴリズムと衛星データの選定

### 4.1 背景と目的

前章までに述べたように、SAR データを用いた作付作物の分類では、高精度化や少ないデータ数での効率化を図ることが求められている。SAR 画像を用いた作付作物の分類手法には、機械学習アルゴリズムを利用した教師付き分類が利用されつつある。その中でも、複数の決定木を用いた分類手法である Random Forests (RF) (Breiman, 2001) は、パターン認識の手法を用いて高い精度が期待できるアルゴリズムの 1 つであり、前章までの結果以外にもいくつかの研究で高精度かつ効率的な作付作物の分類を実行するために寄与している。とくに、サポートベクターマシンなどの分類アルゴリズムと比較して高い精度を示す報告があるだけでなく (Son *et al.*, 2018; Sicre *et al.*, 2020)、ノイズおよび外れ値への高い耐性があることから、複数の研究で衛星データによる高精度な作付作物の分類に寄与している (Zhong *et al.*, 2014; Zhang *et al.*, 2019)。一方、RF のように、複数の決定木を使用したアルゴリズムは他にも開発されており、例えば Extremely Randomized Trees (ERT) (Geurts *et al.*, 2006) および Random Ferns (RFN) (Ozuysal *et al.*, 2010) などでは高い分類の精度を得られた報告がある。これまで ERT は脳腫瘍の自動検出 (Soltaninejad *et al.*, 2017)、RFN はチームスポーツにおける選手の自動検出 (Parisot and Vleeschouwer, 2017) に利用されており、高い精度が得られている。とくに ERT では、光学センサによる作付作物分類でも有用な結果を示した報告がなされている (Kim *et al.*, 2019)。しかし、SAR データによる作付作物分類で ERT や RFN が使用された例はなく、RF との精度面の違いが不明である。

作付作物の分類を行うにあたっては、様々な波長の SAR データが用いられている。先行研究では、多時期の X バンド SAR データを使用することにより、高い精度での作物分類の結果が報告されているほか (Sonobe, 2019)、多時期の C バンド SAR データによる作付作物の分類に関する研究でも高い精度が得られたとの報告がみられる (Skakun *et al.*, 2015; Xie *et al.*, 2019)。また、1~2 時期の C バンドと X バンドのデータを併用することにより、相互に補完し合うことで精度が上昇することも示されている (Zhao *et al.*, 2014)。このように、X バンド SAR や C バンド SAR による作付作物の分類、および両者を併用することによる有効

性を示した研究はいくつかみられるものの、Xバンド SAR と Cバンド SAR の分類の精度を直接比較した研究や、分類に使用する衛星データ数の削減を検討した研究は少ない。少ないデータ数での効率化を図るためには、分類に使用するデータを取捨選択する研究を行う必要がある。

これらの先行研究では、SAR データから主要な数値である後方散乱係数のみを算出し使用している。一方、後方散乱係数だけでなく、各種の散乱成分分解手法から算出される散乱成分を併用することにより、作付作物分類や生育状況推定を行ううえで精度が向上した報告もなされている (Xu *et al.*, 2014)。前章では、Cバンドの RADARSAT-2 データを用いた RF による作付作物の分類で、3 時期以上を使用することにより 0.9 を超える OA が得られた。とくに、RADARSAT-2 データの後方散乱係数、van Zyl *et al.* (2011) による散乱モデル分解法の散乱成分、Cloude and Pottier (1996) による固有値解析法の散乱成分の 3 データセットを組み合わせることにより、最も高い精度で分類を行う事ができた。また、RADARSAT-2 データのこれら 3 データセットに、Xバンドの TerraSAR-X データを併用することにより、2 時期以上を使用することで 0.9 を超える OA が得られた。しかし、ここで使用した 2 偏波の TerraSAR-X データからは、後方散乱係数と Cloude (2007) による固有値解析法の散乱成分の 2 データセットのみを算出しており、2 偏波 SAR の散乱モデル分解法は前章において使用していない。Raney (2013) は、M-chi Decomposition で 2 偏波のための散乱モデル分解法を提案した。この手法を用いることにより、2 偏波のデータでも後方散乱係数、散乱モデル分解法の散乱成分、固有値解析法の散乱成分の 3 データセットを組み合わせることができるようになる。このように、Cバンド SAR および Xバンド SAR の双方から 3 データセットを算出することができるため、これらを用いることでより高精度かつ効率的な分類を行うことができるようになりうると考えられる。

本章では、作物の生育時期に観測された 5~6 時期の Cバンドおよび Xバンドの SAR データを併用した畑作地帯における作付作物の分類を行い、機械学習アルゴリズム間やバンド間での精度の比較や、双方のバンドの一部を組み合わせる手法を検討し、高精度かつ効率的な分類手法の提示を目的とした。SAR データからは、3 章で良好な精度が確認された後方散乱係数と各種散乱成分を分類処理の変数として使用した。まず、分類アルゴリズムとして、RF, ERT, RFN の計 3 種類を使用し、精度の比較を行った。とくに、教師データの比率や衛星データの時期数を変更し、SAR による作付作物分類に適した分類アルゴリズムを選定

した。次に、CバンドとXバンドの各 SAR データでそれぞれ作付作物分類を実行し、各条件下で分類の精度を比較した。実利用の観点からは、できるだけ少ない時期かつ早い時期の SAR データで高い精度を得る必要がある。先述のとおり、分類に使用する衛星データ数の削減を検討した研究は少なく、時期ごとの精度を C バンドと X バンドで比較した研究はみられない。そこで、本章では作付作物分類における各バンドの分類結果や有効な変数について分析し、作付作物の分類に適したバンドを選定した。最後に、その結果を踏まえ、使用する衛星データの利用コストを減少させるために、各バンドの SAR データをすべての時期は使用せずに、効率的に高い精度が得られる衛星データの選定方法を提示した。

## 4.2 研究方法

### 4.2.1 使用データおよび処理方法

本章の対象地域および対象作物は 2 章、3 章と同じである。使用した衛星データは、3 章と同様に C バンドのセンサを搭載した RADARSAT-2 の 4 偏波データおよび X バンドのセンサを搭載した TerraSAR-X の 2 偏波データとした。両衛星の取得した時期は、3 章と同じ 2015 年の 5 時期のほか、2016 年の 6 時期を取得し、2 か年にわたって調査解析を実施した。各衛星データの諸元は 3 章と同じであるが、2016 年のデータは 2015 年と観測のタイミングを変えるため、異なる入射角のデータを使用した。表 4-1 に、衛星データの観測時期および入射角を示す。以降、それぞれの観測時期を早い方から 1 期～5 期、1～6 期と記述する。

3 章で述べたように、2015 年は、TerraSAR-X および RADARSAT-2 を 5 時期ずつ取得しており、それぞれ近い時期のデータを同一の入射角で取得した。一方で 2016 年は、TerraSAR-X および RADARSAT-2 を 6 時期ずつ取得した。2015 年は RADARSAT-2 で 1 種類の入射角のデータしか使用していないため、各観測時期の間隔が 24 日間と開いていた。2016 年は RADARSAT-2 における観測時期の間隔を狭め、作物をより短い間隔で把握するため、6 月中旬から 9 月上旬までの異なる 2 種類の入射角で観測された画像を取得した。作物の判別に寄与すると考えられる時期のデータをより短い間隔で取得することにより、精度の向上も期待できる。TerraSAR-X は、中間配置を目的として、各 RADARSAT-2 の観測時期の中間に観測された画像を取得した。

表 4-1 SAR データの観測時期と公称の入射角

年	時期	RADARSAT-2		TerraSAR-X	
		撮像日	入射角	撮像日	入射角
2015年	1期	05月26日		05月22日	
	2期	06月19日		06月24日	
	3期	07月13日	39.2° ~ 40.7°	07月16日	41.7° ~ 42.9°
	4期	08月06日		08月07日	
	5期	08月30日		08月29日	
2016年	1期	06月13日	39.2° ~ 40.7°	06月21日	
	2期	06月27日	29.1° ~ 30.9°	07月02日	
	3期	07月07日	39.2° ~ 40.7°	07月13日	41.7° ~ 42.9°
	4期	07月21日	29.1° ~ 30.9°	08月04日	
	5期	08月14日	29.1° ~ 30.9°	08月26日	
	6期	09月07日	39.2° ~ 40.7°	09月17日	

表 4-2 算出した各データセットの要素

衛星	データセット	算出した要素
RADARSAT-2	後方散乱係数	HH偏波, HV偏波, VV偏波
	Cloude-Pottier法散乱成分 (CP)	エントロピ (H), アニソトロピ (A), アルファ角 ( $\alpha$ )
	Van Zyl法散乱成分 (VZ)	2回散乱 (Pd), 体積散乱 (Pv), 表面散乱 (Ps)
TerraSAR-X	後方散乱係数	HH偏波, VV偏波
	Cloude法散乱成分 (CD)	エントロピ (H), アルファ角 ( $\alpha$ )
	Raney法散乱成分 (RN)	2回散乱 (Pd), 体積散乱 (Pv), 表面散乱 (Ps)

このように衛星観測のタイミングを決定し、取得されたデータを用いて、分類に使用する変数を算出した。RADARSAT-2 については、前章で有効な変数であった各偏波の後方散乱係数、Van Zyl による散乱モデル分解法の散乱成分 (VZ) (van Zyl *et al.*, 2011)、Cloude-Pottier による固有値解析法の散乱成分 (CP) (Cloude and Pottier, 1996) を算出した。以降では、後方散乱係数や各散乱成分をデータセットと呼ぶ。算出した各データセットの要素を表 4-2 に示す。以上の 3 データセットの計 9 要素を 2015 年は 5 時期、2016 年は 6 時期すべてのデータから算出した。

TerraSAR-X についても、前章で有効であった各偏波の後方散乱係数および Cloude 法による固有値解析法の散乱成分 (CD) (Cloude, 2007) を算出した。また、その他に、Raney によ

る M-chi Decomposition (Raney, 2013) を用いて散乱モデル分解法の散乱成分 (RN) を算出した。以上の 3 データセットの計 7 要素を 2015 年は 5 時期, 2016 年は 6 時期すべてのデータから算出した。

作付作物分類を実施するにあたり, 前章と同様にスペックルフィルタやマルチルック処理, オルソ補正を行った。ただし, Raney の散乱成分分解に関しては, SNAP に処理機能が搭載されていない。そのため, 同じ ESA によって頒布されている PolSARpro SAR Data Processing Educational Tool (PolSARpro) (Pottier and Ferro-Famil, 2012) を使用して値を算出することにした。ただし, PolSARpro ではオルソ補正が行えないため, Alaska Satellite Facility によって頒布されている MapReady Remote Sensing Toolkit (MapReady) (Gens *et al.*, 2013) を使用し, オルソ補正を行った。

#### 4.2.2 解析方法

本章でも 2 章と同様の方法で, 対象地域において十勝農業共済組合より提供を受けた 2 か年の圃場情報 GIS データを用いて, バッファ処理や平均値の抽出を行った。分類処理も同様に, ホールドアウト法により, GIS データに含まれる約 5,000 の圃場ポリゴンのうち 20% を無作為に抽出し, 分類処理のための教師データとした。残りの 80% の圃場ポリゴンはテストデータとし, このテストデータに対する分類結果を基に精度を評価した。ただし, 教師データについては, 実利用時における取得コスト削減の観点から, できるだけ少ないことが望ましい。そこで本章では, この教師データとテストデータの比率 (20% と 80%) を基準とし, 一部で比率を変化させた解析も行った。教師データとテストデータに用いた作物種別の圃場数を表 4-3 に示す。

また, 本章では Random Forests (RF) のほか, その派生形のアルゴリズムである 2 種類の機械学習アルゴリズムを使用して分類の比較を行った。ここで使用した分類手法は, Random Forests のほか Extremely Randomized Trees (ERT) (Geurts *et al.*, 2006) および Random Ferns (RFN) (Ozuysal *et al.*, 2010) である。ERT や RFN は, SAR を利用した作付作物の分類に使用された例がなく, RF と比較した精度が不明であり, 利用価値を検討する必要がある。これらのアルゴリズムを使用するにあたり, フリー統計ソフトである R と randomForest パッケージのほか extraTrees パッケージ (Simm *et al.*, 2014) と rFerns パッケージ (Kursa, 2014) を利用した。

表 4-3 各作物種別の圃場数の内訳

作物	2015年			2016年		
	教師データ	テストデータ	合計	教師データ	テストデータ	合計
小麦	287	1,147	1,434	282	1,135	1,417
豆類	209	837	1,046	190	758	948
馬鈴薯	149	597	746	142	567	709
甜菜	112	446	558	112	447	559
牧草	131	522	653	126	505	631
トウモロコシ	128	511	639	118	474	592
合計	1,016	4,060	5,076	970	3,886	4,856

3 種類のアルゴリズムでは、まず教師データから複数の分類モデル(決定木)を作成する。本章では、各アルゴリズムの決定木の数はそれぞれ 500 とした。次に、作成されたすべての決定木にテストデータを投入する。各決定木では分類結果がそれぞれ導出され、これらの結果の多数決により最終的な分類結果が決定する。

ERT では、決定木を作成する際に、教師データにブートストラップサンプリング法を使用しない。そのため、一部のデータセットではなく全ての教師データがデータセットとして使用される。また、各ノードで使用する変数はランダムに選択され、決定木の分岐関数はジニ係数減少量かエントロピ減少量のいずれかを無作為に選択し作成され、最終的な決定木が作成される (Geurts *et al.*, 2006)。

RFN では、RF と同様に、教師データに対してブートストラップサンプリング法を用いて複数のデータセットの作成を行う。これらのデータセットのうち 3 分の 2 を使用して、決定木を作成する。決定木の分岐関数は、RF と同様にジニ係数減少量を使用して作成される。ただし RF と異なり、それぞれの決定木内において同じ深さのノードでは同じ変数を元とした同じ分岐関数が使用され、決定木が作成される (Ozuysal *et al.*, 2010)。

本章でも、分類の精度の評価には全体精度 (OA)、カッパ係数 ( $\kappa$ )、プロデューサーズ精度 (PA)、ユーザーズ精度 (UA) の 4 種類を使用した。また RF では、各変数におけるジニ係数減少量を算出することで、分類における変数ごとの重要度を示した。ジニ係数の減少量が高い変数ほど分類に有益であることになるため、本章でも分類に寄与した変数を検討するためにジニ係数減少量を使用した。2つの手法間における  $\kappa$  の差の有意性については、検定統計量 Z を使用して評価した。また本章では、2つの分類結果における差の有意性について

て、McNemar 検定 ( $\chi^2$ ) (McNemar, 1947) も使用して評価した。 $\chi^2$  値が 3.86 を超えれば、5% の有意水準において、2 つの分類結果に有意差があると判断することができる。

#### 4.2.3 解析の流れ

本章では、以下の 1), 2) で示す 2 処理で解析を進めた。

##### 1) 作付作物分類における各種機械学習アルゴリズムの精度比較

3 種類の決定木を用いた機械学習アルゴリズム (RF, ERT, RFN) を使用して分類を行い、それぞれの精度を比較した。使用したデータは、2015 年に観測された RADARSAT-2 データおよび TerraSAR-X データである (5 時期, 2 衛星, 計 80 変数)。これらの分類結果から、高い精度が得られたアルゴリズムの特定や、作物ごとの精度の傾向の違いを比較した。

次に、教師データの割合を変化させて精度の比較を行った。ここまでは全て同じ教師データの割合 (全体の 20%) で分類を行ってきたが、教師データの割合をより多くした場合や少なくした場合でも分類を行った。この結果から、3 種類のアルゴリズムにおいて、教師データの割合を変更した場合での精度の違いを比較した。

最後に、使用する時期数を限定して分類を行った。5 時期すべての衛星データを使用しなくても同等の高精度な分類が可能であるかどうかを、アルゴリズムごとに比較した。これらの結果から、アルゴリズムによって OA や作物ごとにおける精度の特徴を比較した。

##### 2) RADARSAT-2 データと TerraSAR-X データとの分類結果の比較

異なるバンドの衛星データで同じ種類のデータセット (後方散乱係数, 散乱モデル分解法, 固有値解析法) を使用した分類において、OA や作物ごとの精度を比較し、分類に用いるための最適なデータについて考察した。使用したデータは、2015 年および 2016 年の RADARSAT-2 データおよび TerraSAR-X データであり、分類には RF を使用した。

まず、1 時期を使用した分類において、RADARSAT-2 データを使用した場合と TerraSAR-X データを使用した場合で精度の差を比較した。バンド間での比較だけでなく、観測された期ごとの精度も比較することで、1 時期での分類に最適なバンドと時期の組み合わせを考察した。

次に、多時期の SAR データを組み合わせた場合の分類を行った。実利用の観点から、現場においてはできるだけ早い時期に分類に使用するデータを取得する必要がある。本章ではこの点を鑑み、1 期から逐次時期を追加していき、1 期のみを使用した場合から 2015 年

は5期まで、2016年は6期までの全てを使用した場合まで、各時期数で分類を行った。その結果から、RADARSAT-2 および TerraSAR-X において、期を追加していった場合での精度の変化の傾向を比較した。RF では、各変数のジニ係数減少量を算出することで、分類における変数ごとの重要度を導出することができる。ジニ係数減少量が高い変数ほど分類に有益であったことを示しているため、この値から各変数の重要度を比較した。

最後に、これまで得られた傾向を踏まえて、データの取得コストを下げることを想定し、できるだけ少ない画像数で効率よく高精度な分類を行う手法を検討した。分類アルゴリズムは RF および ERT を使用し、衛星データは時期数を固定し異なるバンドを併用して分類を行った。使用する衛星データ数を2015年は5時期、2016年は6時期に固定し、各期でいずれかのバンドの衛星データのみを使用し、限られた時期数で高い精度を得る方法を検討した。

## 4.3 結果と考察

### 4.3.1 機械学習アルゴリズムの比較

2015年に取得された5時期のRADARSAT-2データおよびTerraSAR-Xデータを併用し、3種類の分類アルゴリズムで作付作物の分類を行い、それぞれの精度を比較した。3章で最も良好な精度が得られた変数と同じ組み合わせ、すなわちRADARSAT-2データから算出される3データセット(後方散乱係数, CP, VZ)の9要素・5時期(45変数)と、TerraSAR-Xデータから算出される2データセット(後方散乱係数, CD)の4要素・5時期(20変数)を使用した(計65変数)。分類手法は、3種類(RF: Random Forests, ERT: Extremely Randomized Trees, RFN: Random Ferns)のアルゴリズムを使用した。

#### 1) 各分類手法の比較

各機械学習アルゴリズムによる分類結果を表4-4に示す。3種類の分類アルゴリズムの中で最も高い精度となったのはERTであり、OAは0.938、 $\kappa$ は0.923となった。次に高い精度であった分類アルゴリズムはRFであり、OAは0.934、 $\kappa$ は0.919となった。したがって、ERTのアルゴリズムを使用することで、RFよりも高い精度での作付作物分類が行えることが示された。ERTとRFの間で、 $\kappa$ に有意差は生じなかったが、McNemar検定では有意差が

表 4-4 2015 年の RADARSAT-2 および TerraSAR-X データを使用した

作付作物分類における分類アルゴリズム間の結果比較

精度	作物	RF	ERT	RFN
プロデュー サーズ 精度 (PA)	豆類	0.932	0.929	0.879
	甜菜	0.919	0.944	0.964
	牧草	0.952	0.946	0.948
	トウモロコシ	0.835	0.839	0.823
	馬鈴薯	0.926	0.934	0.911
	小麦	0.981	0.982	0.957
ユーザーズ 精度 (UA)	豆類	0.917	0.934	0.951
	甜菜	0.960	0.957	0.870
	牧草	0.925	0.935	0.879
	トウモロコシ	0.896	0.894	0.817
	馬鈴薯	0.899	0.903	0.908
	小麦	0.975	0.970	0.983
全体精度 (OA)		0.934	0.938	0.917
カッパ係数( $\kappa$ )		0.919	0.923	0.898

RF: Random Forests, ERT: Extremely Randomized Trees, RFN: Random Ferns

あった。McNemar 検定で有意差が生じた理由として、RF はブートストラップサンプルから決定木を作成していき、ジニ係数減少量を用いて分岐関数を作成するが (Breiman, 2001), ERT はサンプル全体から決定木が作成され、分岐関数の決定も無作為に 2 種類の手法から選ばれることが考えられる (Geurts *et al.*, 2006)。このように、同じ決定木を用いたアルゴリズムでありながら、決定木の作成方法が異なるため、分類結果に差が生じやすかったと考えられる。とくに甜菜の PA や豆類の UA で精度の向上がみられたことから、ランダムな変数選択や両方の分岐関数の使用による多様な決定木での多数決が、甜菜の正確な判別や他の作物の豆類への誤分類の抑制に有用であることが示唆される。

一方、RFN については、OA が 0.917、 $\kappa$  が 0.898 と、RF や ERT よりも低い精度となった。ERT および RF との  $\kappa$  には有意差があり、McNemar 検定においても有意差がみられた。したがって、本章における作付作物分類では、RFN は ERT や RF と同じ水準の精度で分類を行うことができなかった。RFN で高い精度が得られなかった原因として、同じノードに同じ分岐関数が使われる RFN では (Ozuysal *et al.*, 2010), ERT や RF と比較して柔軟性に富ん

だ多様な決定木が作成されなかったことが考えられる。RFN では、甜菜、牧草、トウモロコシの UA が特に低下しており、他の作物をこれらの作物に誤分類する傾向がみられた。したがって、これらの作物は、分岐が進んだ後のノードにおいて、それぞれ別々の分岐関数を使用することにより、更に正確な判別ができると考えられる。そのため、RF や ERT では高い UA で分類がなされたと考えられる。

作物ごとの精度は、各アルゴリズムで全体的に同様の結果となった。小麦は最も高い精度で分類が可能であったが、トウモロコシは豆類や馬鈴薯に誤分類され、高い精度が得られなかった。これらは 3 章でみられた傾向とも一致している。

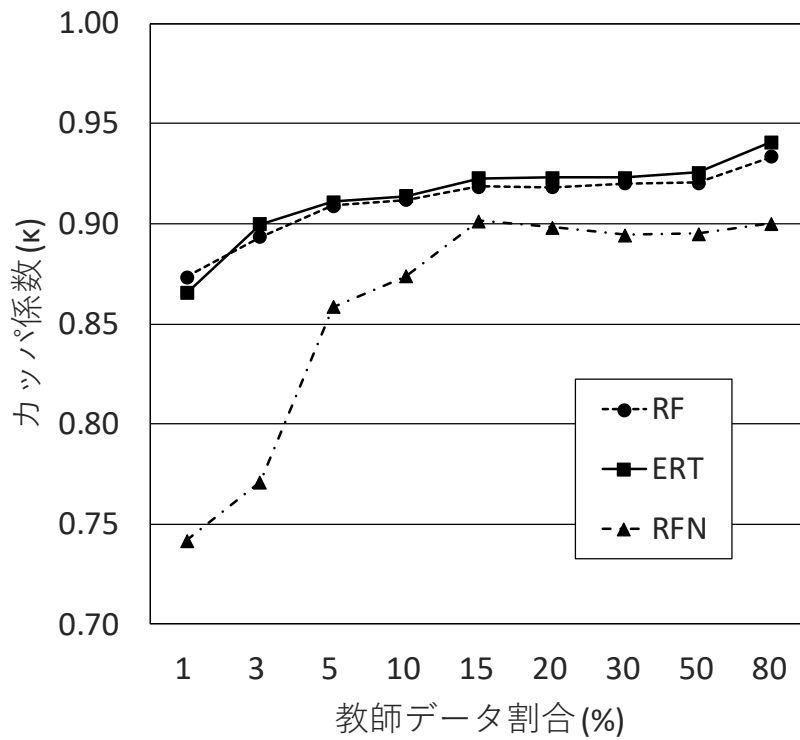
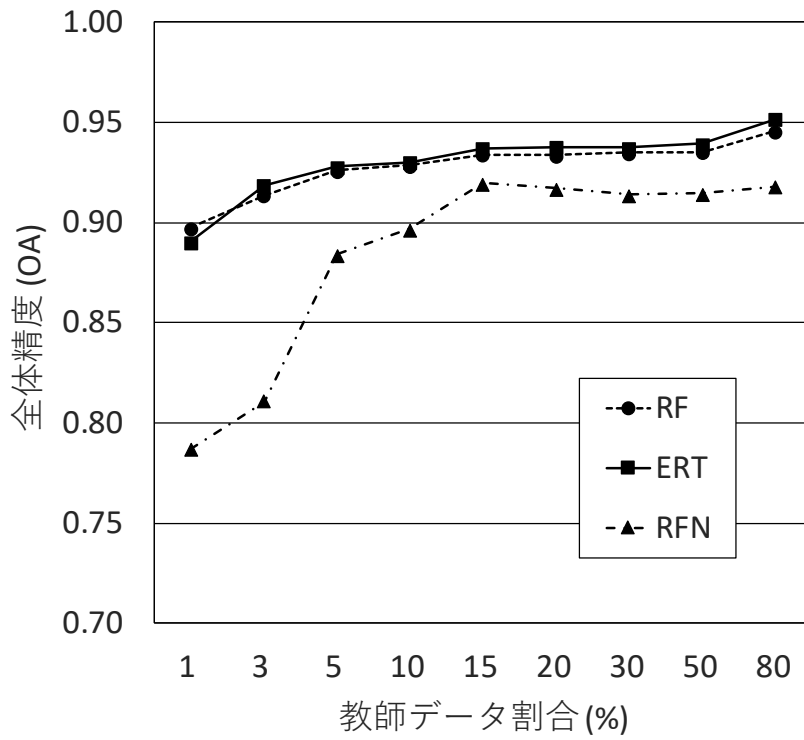
## 2) 教師データの割合を変更した各分類手法の比較

つづいて、これら 3 種類の機械学習アルゴリズムについて、教師データの割合を変化させた場合の精度を比較した。作成した教師データの割合は、その他の解析で使用した 20%のほか、より多い 30%、50%、80%、より少ない 15%、10%、5%、3%、1%の 9 パターンである。教師データの割合とその圃場数を表 4-5 に示し、使用した教師データの割合に対する各アルゴリズムの OA および  $\kappa$  を図 4-1 に示す。

分類アルゴリズム間で同じ教師データの割合ごとに OA および  $\kappa$  を比較すると、1) と同

表 4-5 2015 年 RADARSAT-2 データおよび TerraSAR-X データを使用した  
作付作物分類における各分類アルゴリズムの教師データ割合と圃場数

教師データ 割合 (%)	教師データ 圃場数	テストデータ 圃場数
1	51	5,025
3	152	4,924
5	254	4,822
10	508	4,568
15	761	4,315
20	1,015	4,061
30	1,523	3,553
50	2,538	2,538
80	4,061	1,015



RF: Random Forests, ERT: Extremely Randomized Trees, RFN: Random Ferns

図 4-1 2015 年 RADARSAT-2 データおよび TerraSAR-X データを使用した作付作物分類における各分類アルゴリズムの教師データ割合と精度との関係

様に、ほぼすべての場合で ERT が最も高い OA となり、教師データの割合を 1%にした場合のみ、RF の方が高い OA となった。しかし、ERT と RF との間にはどの教師データ数の割合であっても  $\kappa$  に有意な差はなかった。したがって、ERT では教師データの割合が極めて少ない場合に限り RF よりも精度が低下する可能性はあるものの、ERT と RF は教師データの数に関わらず同等の精度で分類が可能といえる。

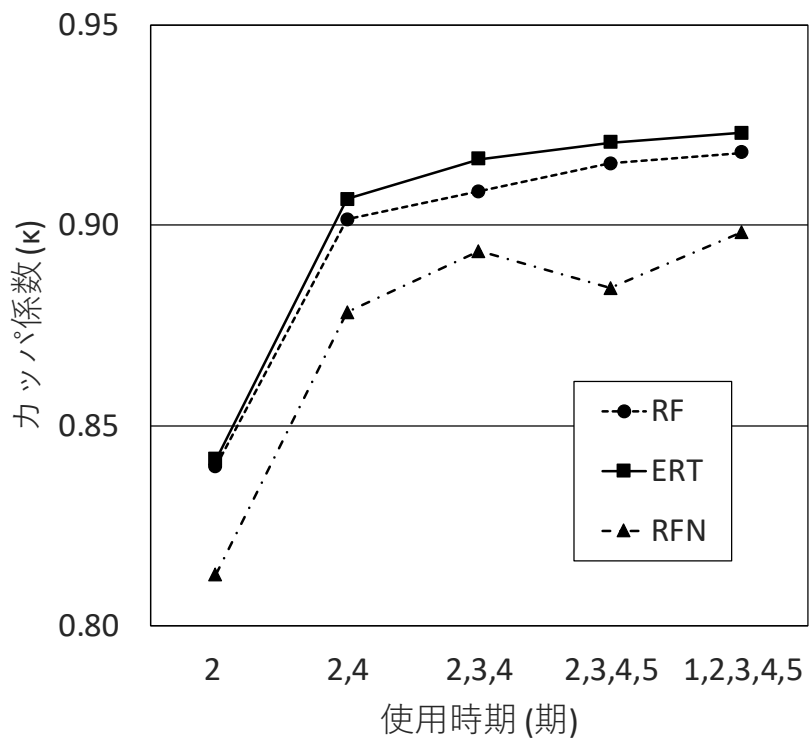
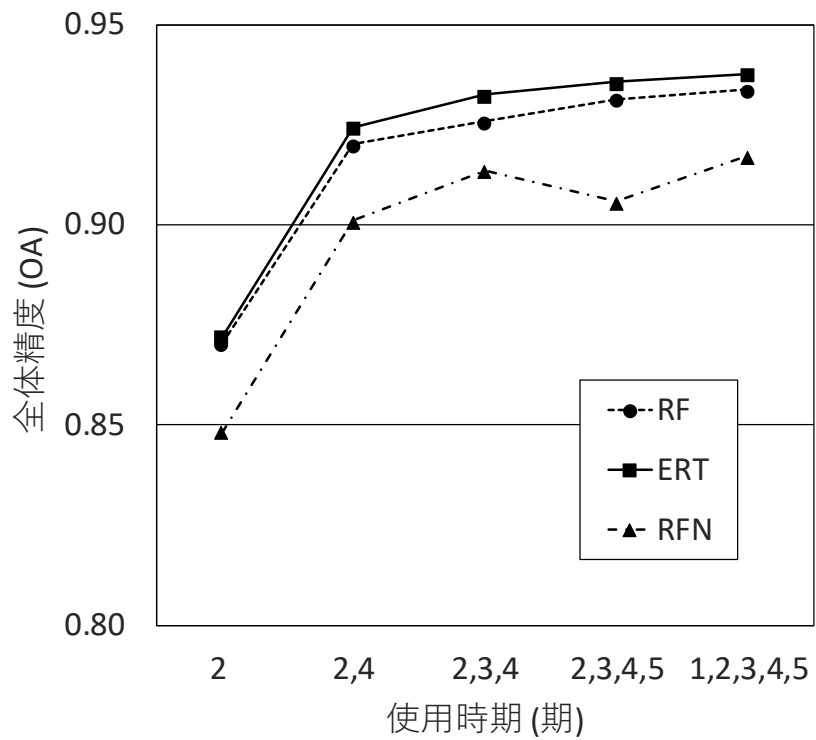
この教師データの割合を変化させた分類においても、RFN はおしなべて最も低い精度となった。RFN の  $\kappa$  を RF や ERT と比較したところ、どの教師データ数であっても有意な差がみられた。この結果から、RFN は教師データ数を変化させても RF や ERT と同等な  $\kappa$  で作付作物分類を行うことは不可能である。すなわち、本章のように使用した変数の数が多い場合には、同じノードで同じ分岐関数を用いる RFN では、RF や ERT ほど高精度な分類が見込めないと考えられる。

教師データの割合を減少させると、いずれのアルゴリズムでも OA や  $\kappa$  は低下するが、ERT と RF は全体的に高い精度が維持された。両者とも、教師データが 3%以上の割合で OA は 0.9 を超えていた。ERT、RF とも、教師データの割合を 80%から 50%に下げると若干の精度低下がみられたが、80%との間に  $\kappa$  の有意な差はなかった。教師データの割合が 15%以上であれば、80%との間に  $\kappa$  の有意な差がなかった。したがって、本章のように 5,000 圃場の作付作物を分類する場合、教師データが 15%以上の割合 (約 750 圃場) であれば十分な精度が得られるといえる。また、5% (約 250 圃場) を下回ると精度低下が大きくなった。

RFN については、ERT および RF と比較して、15%から 80%まで精度の変化が小さく、 $\kappa$  に有意な差もなかった。しかし、教師データの割合が 15%を下回ると、ERT や RF と比較しても更に大幅な精度の低下がみられ、教師データの割合が 1%になると OA も 0.80 を下回る結果となった。したがって、RFN は RF や ERT と比較して、教師データ数によって精度が左右されやすいといえる。

### 3) 使用時期数ごとの各分類手法の比較

これら 3 種類の機械学習アルゴリズムについて、使用する時期数を限定して分類を行い、精度の変化を比較した。なお、各時期数において使用したデータの期は、最も精度が高かった組み合わせとした。1 時期のみでは 2 期、2 時期では 2 期と 4 期、3 時期では 2 期から 4 期、4 時期では 2 期から 5 期とした。使用した期に対する各アルゴリズムの OA および  $\kappa$  を図 4-2 に示す。



RF: Random Forests, ERT: Extremely Randomized Trees, RFN: Random Ferns

図 4-2 2015 年 RADARSAT-2 データおよび TerraSAR-X データを使用した  
作付作物分類における各分類アルゴリズムの使用時期数と精度との関係

分類アルゴリズム間で同じ時期数ごとに OA および  $\kappa$  を比較したところ、こちらもすべての場合で ERT が最も高い OA となった。その差は 3 時期使用した場合に最も大きく、使用する時期数を少なくしたり多くしたりすることにより、ERT と RF との OA の差は小さくなる傾向がみられた。このことから、ERT は RF と比較すると高い精度で分類を行うことができ、使用する時期数が 3 時期程度の場合に有用性が高くなる傾向が示唆された。2) と同様に、同じ時期数ごとに ERT と RF を比較した結果、どの時期数でも  $\kappa$  に有意差はみられなかったが、McNemar 検定では有意な差がみられた。

RFN については、これまでの結果と同様に全ての時期数で最も低い精度となった。また、RFN は RF や ERT との間で  $\kappa$  を比較すると、どの教師データ数であっても有意な差があった。したがって、RFN は教師データ数を変化させても RF や ERT と同等な水準の高い精度で作付作物分類を行うことは不可能である。

各アルゴリズムで時期ごとに精度の比較をしたところ、3 種類のアルゴリズムすべてにおいて、2 時期以上を使用すると OA は 0.9 を上回り、3 時期と 5 時期で  $\kappa$  に有意差がなかった。また、RF による 5 時期のデータを使用した分類と、ERT における 2 時期のデータを使用した分類にも  $\kappa$  に有意差がなかった。ERT は、2 時期のデータでの分類でも RF での 5 時期での分類と同等の精度であり、RF と比較して少ない時期数においても効率的に高い精度が得られたといえる。RFN では、4 時期の使用にて 3 時期を使用した際より精度の低下がみられた。この原因として、同一ノードに同一の分岐関数しか使われないために、過学習が発生した可能性が考えられる。

#### 4.3.2 RADARSAT-2 データと TerraSAR-X データとの分類結果の比較

C バンドである RADARSAT-2 データから算出された 3 データセット（後方散乱係数, VZ, CP) の 9 要素と、X バンドである TerraSAR-X データから算出された 3 データセット（後方散乱係数, CD, RN) の 7 要素を用いて (表 4-2), X バンドおよび C バンドにおける分類の精度の差異やその特徴を、様々な条件下で比較した。使用した衛星データは、2015 年の 5 時期 (計 80 変数) および 2016 年の 6 時期 (計 96 変数) のデータである。分類の手法には、RF のほか、4.3.1 において高い精度での分類が可能となった ERT も使用した。なお、本節のみ、以降 RADARSAT-2 を C バンド、TerraSAR-X を X バンドと呼ぶ。

### 1) 1 時期の SAR データのみを使用した分類

まず、各バンドで 1 時期のみの SAR データを使用して分類を行った。使用した 3 手法のうち ERT と並んで高い精度が得られた RF の結果を表 4-6 に示す。各年の各期において C バンドと X バンドの結果を比較し、OA と  $\kappa$  が高かった方のバンドの結果を太字で示した。2015 年において、1 期、3 期では C バンドの方が高い精度で分類がなされ、2 期、4 期、5 期では X バンドの方が高い精度で分類がなされた。ただし、3 期については両者の  $\kappa$  に有意な差がなかった。2016 年もほぼ同様の傾向がみられ、1 期、3 期、6 期では C バンドの方が高い精度で分類がなされ、2 期、4 期、5 期では X バンドの方が高い精度で分類がなされた。ただし、1 期については両者の  $\kappa$  に有意な差がなかった。したがって、各年において、7 月までは C バンドの方が概ね高い精度を得ることができる傾向にあり、8 月中は X バンドの方が高い精度を得ることができる傾向にあることが示された。すなわち、1 時期のみの SAR データで分類を行う場合、生育シーズンの初期における作付作物は C バンドで、生育シーズンの後期における作付作物は X バンドで、それぞれ分類を行うことで高い精度が得られることが示唆された。

この理由として、X バンドと C バンドで地表面の物体における透過の度合いが異なっていることが考えられる。X バンドは C バンドと比較して波長が短く、作物体を透過しにくい傾向にある。したがって、作付作物の生育が進み、作物の種類ごとに草型に差異が生じてくるシーズン後期において、X バンドは作物を詳細に判別することができ、高い精度が得ら

表 4-6 1 時期の SAR データのみを使用した RF による分類における精度の比較

使用時期	2015年				2016年					
	RADARSAT-2 (Cバンド)		TerraSAR-X (Xバンド)		RADARSAT-2 (Cバンド)		TerraSAR-X (Xバンド)			
	OA	$\kappa$	OA	$\kappa$	OA	$\kappa$	OA	$\kappa$		
1期	<b>0.629</b>	<b>0.540</b>	0.573	0.465	*	<b>0.806</b>	<b>0.759</b>	0.791	0.741	
2期	0.772	0.719	<b>0.810</b>	<b>0.766</b>	*	0.733	0.669	<b>0.828</b>	<b>0.787</b>	*
3期	<b>0.768</b>	<b>0.713</b>	0.760	0.704		<b>0.853</b>	<b>0.818</b>	0.807	0.761	*
4期	0.689	0.616	<b>0.777</b>	<b>0.725</b>	*	0.748	0.689	<b>0.797</b>	<b>0.749</b>	*
5期	0.698	0.622	<b>0.718</b>	<b>0.648</b>	*	0.703	0.631	<b>0.762</b>	<b>0.706</b>	*
6期						<b>0.729</b>	<b>0.662</b>	0.613	0.513	*

\* $p < 0.05$ , OA: 全体精度,  $\kappa$ : カップ係数  
太字は各年の各期で高い方のバンドのOAと $\kappa$

れたと考えられる。一方、波長の長い C バンドは、比較的作物体を透過しやすい傾向にある。この特徴が、作付作物の生育が十分に進んでいない生育シーズン前期において、畝の形状や土壌表面のわずかな差異の判別や、既に生育が進んでいる小麦や牧草の判別に寄与したと考えられる。2 期に関しては、生育ステージ前期ではあったが X バンドの方が C バンドと比較して高い精度が得られた。作物別に精度を確認したところ、両年ともトウモロコシで C バンドによる UA, PA が著しく低下していた。6 月下旬のトウモロコシはまだ作物体としては小さかったために、他にこの時期で作物体が小さかった豆類と区別がしにくく、誤分類が生じたと考えられる。一方、この時期のトウモロコシは豆類と比較して葉面積は大きいため、X バンドではその特徴を捉え、比較的高い精度で分類がなされていたと考えられる。

各バンドの精度の変化を時期別に見た場合、2015 年の 1 期と 2016 年の 6 期において、OA が極端に低下していた。この理由として、2015 年は 1 期を 5 月下旬に、2016 年は 6 期を 9 月上旬に取得している点があげられる。5 月下旬は小麦や牧草以外の作物で極端に草高が小さく、小豆など出芽をしていない作物もある。また、9 月上旬は小麦や馬鈴薯、トウモロコシなどで既に収穫が終了しており、地上部に植物体が存在しない圃場が多い。したがって、これらの時期のデータではすべての作付作物を十分に分類できるとは限らず、精度が低下しやすいと考えられる。

## 2) 多時期の SAR データを組み合わせた分類

つづいて、複数の時期の SAR データを組み合わせて分類を行った。本章では、1 期から逐次時期を追加していく方法を取ることとし、それらの各分類結果から精度を比較した。また、各バンドにおける精度の変化やその傾向を比較した。本節でも、使用した 3 手法のうち ERT と並んで高い精度が得られ、かつ分類に使用した変数の重要度を算出できる RF による分類を行った。

結果を表 4-7 に示す。各年の各時期数において C バンドと X バンドの結果を比較し、OA と  $\kappa$  が高かった方のバンドの結果を太字で示した。2015 年、2016 年いずれの結果でも、ほぼすべての時期数で C バンドの方が高い OA を示した。この理由として、1) で示したように、1 期や 3 期などの早い時期のデータで C バンドの方が精度は高く、多時期のデータを使用してもその影響が出ていたためであると考えられる。とくに両年とも、1 期から 3 期までを使用した場合については、X バンドと C バンドの間で  $\kappa$  に有意な差がみられた。したが

表 4-7 多時期の SAR データを使用した RF による分類における精度の比較

使用時期	2015年				2016年			
	RADARSAT-2 (Cバンド)		TerraSAR-X (Xバンド)		RADARSAT-2 (Cバンド)		TerraSAR-X (Xバンド)	
	OA	$\kappa$	OA	$\kappa$	OA	$\kappa$	OA	$\kappa$
1期	<b>0.629</b>	<b>0.540</b>	0.573	0.465	<b>0.806</b>	<b>0.759</b>	0.791	0.741
1期～2期	0.814	0.771	<b>0.835</b>	<b>0.795</b>	0.835	0.796	<b>0.857</b>	<b>0.823</b>
1期～3期	<b>0.875</b>	<b>0.845</b>	0.855	0.820	<b>0.906</b>	<b>0.883</b>	0.888	0.861
1期～4期	<b>0.907</b>	<b>0.886</b>	0.902	0.878	<b>0.925</b>	<b>0.907</b>	0.918	0.899
1期～5期	<b>0.917</b>	<b>0.898</b>	0.909	0.888	<b>0.934</b>	<b>0.918</b>	0.929	0.912
1期～6期					<b>0.935</b>	<b>0.920</b>	0.932	0.916

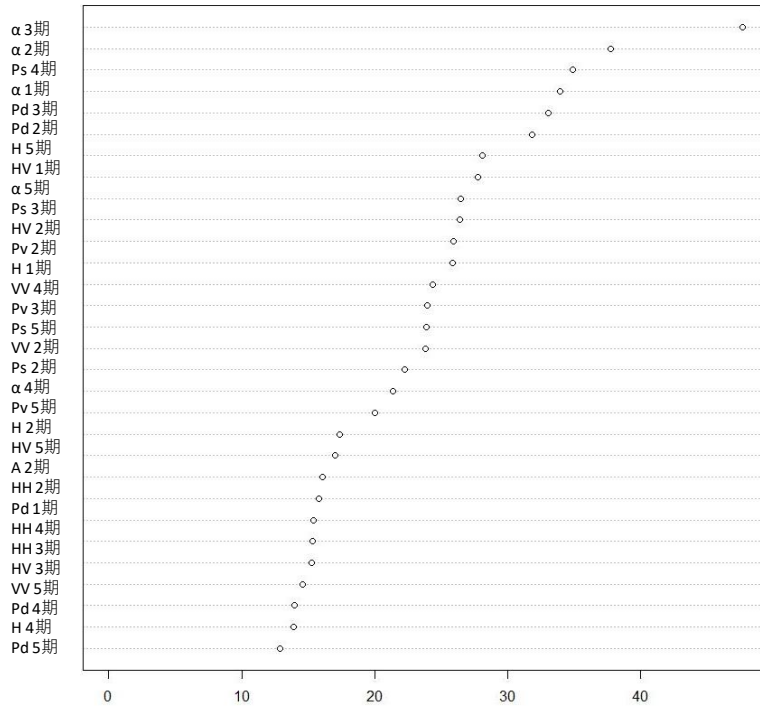
\* $p < 0.05$ , OA: 全体精度,  $\kappa$ : カッパ係数

太字は各年の各時期数で高い方のバンドのOAと $\kappa$

って、7月中までのデータを使用した場合には、Cバンドを使用した分類において高い有用性が示された。4期以降のデータを組み合わせると、両年ともXバンドとCバンドの間で $\kappa$ に有意な差がなかった。すなわち、8月のデータを含めて分類を行った場合には、Cバンドの方が高い精度ではあったものの、XバンドのOAも上昇していったために、両者の間の差は徐々に小さくなった。その結果、いずれのバンドのデータを使用しても $\kappa$ に有意な差がない結果となった。1期から2期までを使用した場合については、いずれの年でもXバンドを使用した分類において高い精度が得られた。これは、1)の2期での分類において、Cバンドデータで高い精度を得られなかったことが要因となっていると考えられる。

すべての時期を使用した分類において、特に重要度の高かった変数を調べるため、各変数のジニ係数減少量を比較した。各バンドにおける時期別のジニ係数減少量について、値の高い(重要度が高い)順に示したものを図4-3、4-4に示す。Cバンドにおいては、いずれの年のデータについても、生育ステージ初期(6月～7月)のアルファ角、表面散乱、2回散乱の各散乱成分が高い重要度を示した。小麦の垂直構造は2回散乱が、作物体の小さい圃場は表面散乱の割合がそれぞれ高く、分類に有効であったと考えられる。そのほか、時期を問わずエントロピ、HV偏波、VV偏波が高い重要度を示した。RADARSAT-2において、エントロピやHV偏波は作物の生長と深い関係があり(Jiao *et al.*, 2014)、VV偏波は植生の有無で値が変化する(Huang *et al.*, 2017)。本章でも作物を判別するために十分な変数となったと考え

### 2015年



### 2016年

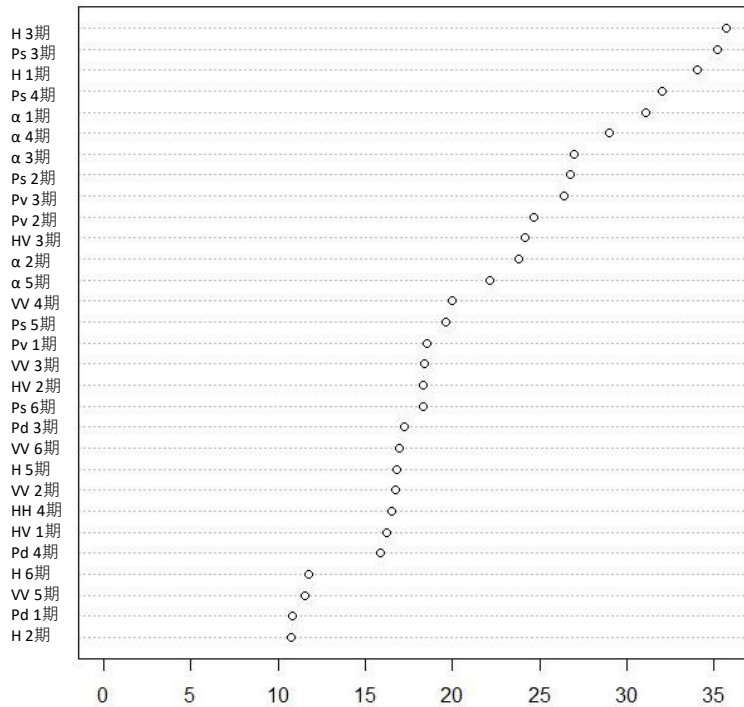
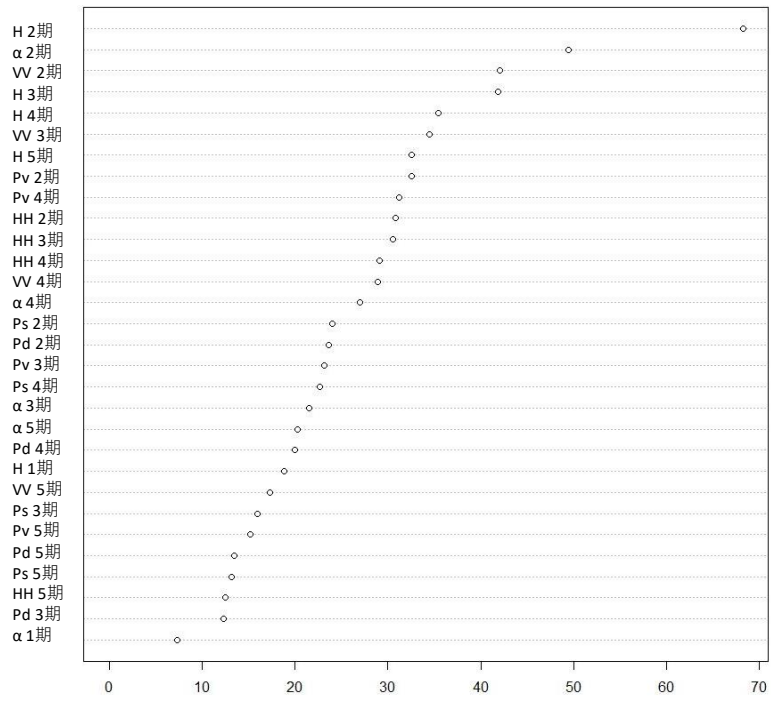


図 4-3 RADARSAT-2 データを使用した分類におけるジニ係数減少量

### 2015年



### 2016年

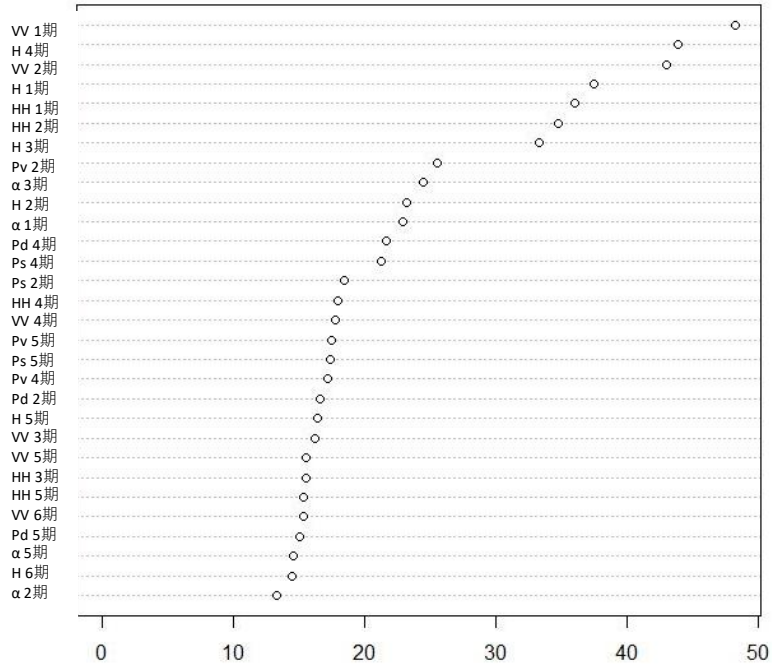


図 4-4 TerraSAR-X データを使用した分類におけるジニ係数減少量

られる。

Xバンドでの分類では、3章の結果と同様、各年とも生育ステージ初期（6月～7月）のアルファ角、HH偏波、VV偏波で高い重要度を示した。その結果に加えて、本章ではエントロピも同様に高い重要度を示した。また、本章で新たに使用した6月上旬や8月のデータも特に重要な役割を果たしていた。さらに、体積散乱も生育ステージの時期を問わず高い重要度を示した。体積散乱は作物のバイオマス量と関係が深いため、作付作物を判別する変数として役割を果たしたと考えられる。

Cバンド、Xバンドとも、5月や9月の変数は高い重要度を得られなかった。作物体が小さいことや、複数の作物が収穫後であることが原因であると考えられる。変数別にみると、HHやVVの後方散乱係数に関しては、CバンドよりXバンドで高い重要度を示していた。さらに、生育ステージ初期の変数で高い重要度を示す傾向はCバンドでより顕著であり、生育ステージ後期の変数が高い重要度を示す傾向はXバンドでより顕著であった。このことは、1)で確認された、生育ステージ前期でCバンドが高い精度を示し、生育ステージ後期でXバンドが高い精度を示した傾向にも結び付くといえる。

### 3) 異なるバンドを併用した分類

できるだけ少ない画像数で効率よく高精度な分類を行うために、使用する時期数を固定して異なるバンドを併用し、高い精度を得る方法を検討した。1) および2)の結果から、Cバンドは生育ステージ前期で精度が高く、Xバンドは生育ステージ後期で精度が高いと考えられた。そこで、生育ステージ前期ではCバンド、生育ステージ後期ではXバンドを使用して分類するという手法を考案した。使用するデータ数を2015年は5時期、2016年は6時期に固定した。分類アルゴリズムには、RFおよびERTをそれぞれ使用した。

2015年、2016年それぞれRFおよびERTで分類を行った結果を表4-8に示す。使用データの列は左から右に1期から5期、1期から6期に使用したバンド（CまたはX）を示している。各年で最も高いOAと $\kappa$ が得られた組み合わせを太字で示した。その結果、いずれの年およびアルゴリズムでも同様の傾向がみられた。

生育初期に使用するデータをCバンドにし、生育後期に使用するデータをXバンドにした分類は、すべて同じバンドを使用した分類と比較して、高い精度が得られた。とくに2016年については、2期までをCバンドにし、3期以降をXバンドにすることで、RF、ERTともに0.95を超える最も高いOAが得られた。さらにこの結果は、すべて同じバンドを使用

表 4-8 異なるバンドの 5, 6 時期の SAR データを併用した分類における

Random Forests (RF) および Extremely Randomized Trees (ERT) の精度の比較

使用データ	2015年				使用データ	2016年			
	RF		ERT			RF		ERT	
1期～5期	OA	$\kappa$	OA	$\kappa$	1期～6期	OA	$\kappa$	OA	$\kappa$
XXXXX	0.909	0.888	0.914	0.894	XXXXXX	0.932	0.916	0.939	0.924
XXXXC	0.913	0.892	0.917	0.898	XXXXXC	0.933	0.918	0.940	0.926
XXXCC	0.906	0.884	0.912	0.891	XXXCC	0.931	0.914	0.933	0.917
XXGCC	0.907	0.886	0.911	0.890	XXGCC	0.925	0.907	0.928	0.911
XCCCC	0.909	0.888	0.915	0.895	XXCCCC	0.930	0.914	0.932	0.916
CCCCC	0.917	0.898	0.921	0.903	XCCCCC	0.935	0.920	0.935	0.920
CCGCX	0.918	0.899	0.924	0.906	CCGCC	0.935	0.920	0.939	0.925
CCGCX	<b>0.924</b>	<b>0.906</b>	<b>0.931</b>	<b>0.915</b>	CCGCX	0.941	0.928	0.943	0.929
CCXXX	0.920	0.901	0.929	0.913	CCGCXX	0.947	0.934	0.951	0.939
CXXXX	0.917	0.898	0.919	0.899	CCGXXX	0.947	0.934	0.949	0.937
					CCXXXX	<b>0.951</b>	<b>0.939</b>	<b>0.953</b>	<b>0.942</b>
					CXXXXX	0.947	0.935	0.950	0.939

RF: Random Forests, ERT: Extremely Randomized Trees, OA: 全体精度,  $\kappa$ : カツパ係数  
 使用データの列は左から右に1期から5期, 1期から6期に使用したバンド(CまたはX)を記載  
 太字は各年で最も高いOAと $\kappa$

した分類と比較して  $\kappa$  に有意差がみられた。2015 年の結果も同様に, 3 期までを C バンド, 4 期以降を X バンドにすることで, RF, ERT とともに 0.92 を超える最も高い OA が得られた。この結果は, C バンドのみを使用した分類と比較すると  $\kappa$  に有意差はなかったが, X バンドのみを使用した分類とは  $\kappa$  に有意差がみられた。一方, 生育初期に使用するデータを X バンドにし, 生育後期に使用するデータを C バンドにした分類は, すべて同じバンドを使用した分類と比較して,  $\kappa$  に有意差はなく精度の向上は確認されなかった。以上より, 同じ時期数でも異なるバンドを併用することで精度を向上させられることが示された。とくに 7 月中旬を境界として, それまでを C バンド, それ以降を X バンドにすることで最も高い精度が得られた。このことは, 1) および 2) の結果と同様であり, 生育ステージ前期の C バンドと生育ステージ後期の X バンドの組み合わせにより高い精度での分類を行えることがここでも示された。2016 年において最も高い精度であった生育ステージ前期の C バンド 2 時期と生育ステージ後期の X バンド 4 時期を使用した ERT の分類結果のマッピングを図 4-5 に示す。

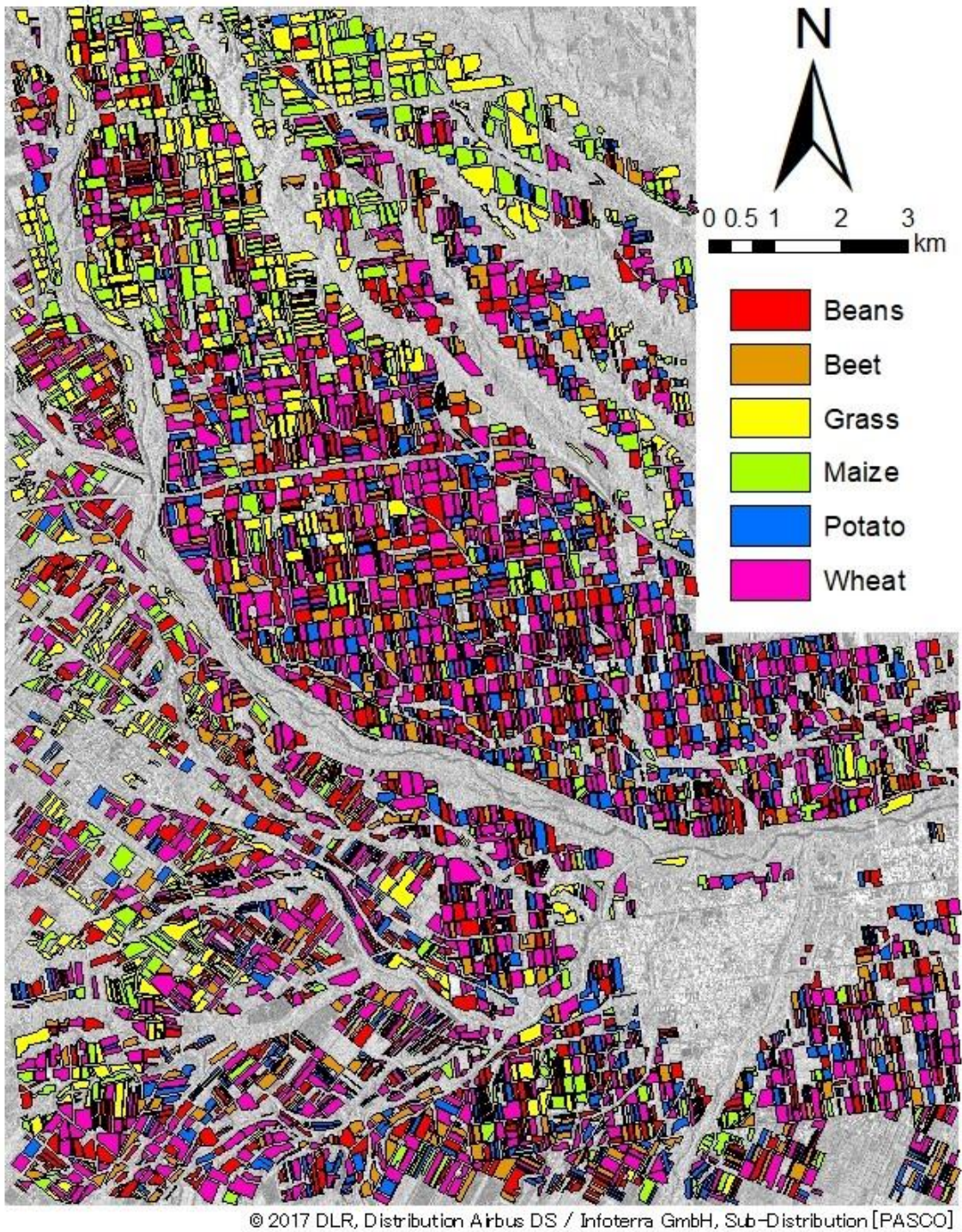


図 4-5 2 期までの RADARSAT-2 データと 3 期以降の TerraSAR-X データを併用した 2016 年の分類マップ (Extremely Randomized Trees による分類)

これらの分類結果について、作物ごとの精度を比較した。2016年のERTにおける分類結果について、最も高い精度であったバンドの組み合わせ(2期までのCバンドと3期以降のXバンドの併用)による分類と、Cバンドのみを使用した分類、Xバンドのみを使用した分類の結果のそれぞれについて、OA、 $\kappa$ 、作物別のPA、UAを表4-9に示す。すべて同じバンドを使用した分類では、Cバンド、Xバンド間で精度の高い作物が異なっていたが、2期までのCバンドと3期以降のXバンドを併用した分類では、各作物のUA、PAが、いずれか高い方のバンドの精度の水準まで向上した。例えば、豆類や甜菜のUAについて見ると、Xバンドで高い値を得ていたが、併用した分類でもその値と同程度のUAが得られた。馬鈴薯のUA、PAについて見ると、Cバンドで高い値を得ていたが、併用した分類でもその値と同程度のUA、PAが得られた。このように、併用した分類では、各バンドの長所を生かすことで、OAが向上していた。トウモロコシのUA、PAについて見ると、Xバンド、Cバンドともに低い値であったが、併用した分類では高い値となった。このように、双方のバンドの変

表 4-9 2期までのRADARSAT-2データと3期以降のTerraSAR-Xデータを併用した2016年の分類におけるERTでの作物別の精度比較

精度	作物	使用データ				
		Cバンド Xバンド	1期～6期 —	— 1期～6期	1期～2期 3期～6期	1期～6期 1期～6期
プロデュー サーズ 精度 (PA)	豆類		0.951	0.942	0.954	0.970
	甜菜		0.915	0.917	0.940	0.928
	牧草		0.917	0.935	0.921	0.939
	トウモロコシ		0.840	0.867	0.916	0.899
	馬鈴薯		0.961	0.937	0.972	0.972
	小麦		0.981	0.978	0.979	0.986
ユーザーズ 精度 (UA)	豆類		0.885	0.930	0.953	0.934
	甜菜		0.930	0.940	0.959	0.965
	牧草		0.963	0.931	0.941	0.969
	トウモロコシ		0.873	0.865	0.906	0.901
	馬鈴薯		0.966	0.938	0.952	0.962
	小麦		0.985	0.979	0.977	0.987
全体精度 (OA)			0.939	0.939	0.953	0.957
カッパ係数( $\kappa$ )			0.925	0.924	0.942	0.947

数で補完し合うことによって、これまで1種類のバンドでは低い精度であった作物も、同じ時期数で高い精度で分類された。

最後に、今回最も高い精度が得られた、2016年のERTにおける7月中旬までのCバンド2時期と、7月中旬以降のXバンド4時期を組み合わせた分類(計6時期)について、Cバンド6時期、Xバンド6時期すべてを使用した分類(計12時期)と精度を比較した。結果を表4-9の右端の列に示す。各バンドを6時期すべてを使用した分類の方が高いOA、 $\kappa$ であったが、Cバンド2時期とXバンド4時期を組み合わせた分類との間に $\kappa$ の有意差はなかった。3章では、Xバンド、Cバンドを併用することにより、どちらか一方のみのバンドの衛星データを使用するよりも高い精度で分類がなされることを示した。しかしながら、この結果から、7月中旬まではCバンドのみ、7月中旬以降はXバンドのみの衛星データを使用するだけで良く、すべての期において双方の衛星データを使用しなくても、同程度の精度で分類が可能であることを示した。すなわち本結果から、使用する衛星データの利用コストにおける大幅な減少を提示することができた。

#### 4.4 まとめ

本章では、高精度な分類とデータ数削減による効率的な分類を、分類アルゴリズムとSARデータのバンドの双方向から検討し、有効な方法を提示した。SARデータとしてRADARSAT-2のCバンド4偏波データと、TerraSAR-XのXバンド2偏波データを使用し、各偏波の後方散乱係数や、散乱モデル分解法および固有値解析法から算出される散乱成分を使用した。分類アルゴリズムとして、決定木による学習アルゴリズムの一種であるRandom Forests (RF) や、Extremely Randomized Trees (ERT)、Random Ferns (RFN) を使用し、教師付き分類を実行した。

まず、決定木を用いた3種類の分類手法について、使用する教師データの比率や時期数の条件を変更し、精度を比較した。その結果、RFNはすべての条件でERTやRFより全体精度(OA)が低く、カッパ係数( $\kappa$ )にも有意差がみられた。ERTは、ほぼすべての条件でRFより高いOAを示したが、教師データを約50まで減少させた場合のみRFの方がERTより高いOAとなった。ただし、ERTとRFの間にはすべての条件で $\kappa$ に有意差はなかった。ま

た、2 時期のデータのみを使用した ERT での分類は、5 時期のデータを使用した RF での分類と  $\kappa$  に有意差がない精度となり、ERT による高精度な分類が少ないデータ数で達成されることを示した。以上の結果から、SAR を用いた多時期の作付作物分類において、ERT が特に有効な分類アルゴリズムであると結論づけた。

次に、この ERT を適用し、X バンド SAR である TerraSAR-X と C バンド SAR である RADARSAT-2 のどちらで高い精度を得られるか、時期や時期数を変更し比較を行った。その結果、1 時期による分類では、7 月までは C バンドの方が概ね高い OA となり、8 月は X バンドの方が高い OA となる傾向がみられた。多時期での分類では、C バンド SAR の方が X バンド SAR と比較してほぼすべて高い OA となり、多時期の作付作物分類に有効なバンドであることを示した。これは、生育ステージ初期における C バンドの変数が分類に寄与したためであると考えられる。以上より、多時期の作付作物の分類をいずれかのバンドで行う場合においては、C バンドの方が有効であると結論づけた。

最後に、これらの結果を踏まえ、使用する衛星データの利用コストを抑制するため、衛星データの選択方法を検討した。その結果、7 月上旬までの C バンドと 7 月中旬以降の X バンドを併用することで、0.95 を超える OA が得られた。これは、両バンドの 6 時期をすべて使用した分類と比較して、 $\kappa$  に有意差がなかった。以上のように、双方のバンドを用いる場合には、7 月中旬で C バンドから X バンドに衛星データを切り替える手法が、衛星データの利用コストを大幅に減少させつつ高い精度を得られるものとして期待される。この手法は、本研究の成果による新たな衛星データの選定法として提示されるものである。



## 5 CバンドおよびLバンド SAR データを併用した機械学習アルゴリズムによる作物分類の効率化に関する検討

### 5.1 背景と目的

Cバンド SAR である RADARSAT-2 データと、Xバンド SAR である TerraSAR-X データを併用した作付作物の分類では、2 時期のデータで 0.9 を超える全体精度を得ることができたほか (3 章)、7 月中旬までの 2 時期の RADARSAT-2 データと 7 月中旬以降の 4 時期の TerraSAR-X データを併用した分類により 0.95 を超える全体精度を得られた (4 章)。これにより、多時期のデータを大量に使用しなくても、作付作物の分類を高い精度で行うことが可能となった。しかし、RADARSAT-2 や TerraSAR-X などの SAR データは基本的に有料であるが、Lバンド SAR である ALOS-2/PALSAR-2 データはその中でも低価格で入手できる (古川・古関, 2019)。このデータを効果的に使用できれば、TerraSAR-X データや RADARSAT-2 データの代替としてコストの削減を図ることができる。また、既に観測した時期については、PALSAR-2 データのアーカイブから必要な時期と観測地域が存在すればデータを購入することができる。

Lバンド SAR の PALSAR-2 の波長帯は約 24cm であり、Cバンド SAR の RADARSAT-2 (3cm) や Xバンド SAR の TerraSAR-X (5cm) と比較して長波長である。波長は長いほど地表面の物体を透過しやすいため、Lバンドはバイオマス量の大きい作物は透過しにくいですが、バイオマス量の小さい作物は透過しやすい。そのため、特にバイオマス量の大きい作物を分類するには PALSAR-2 データの使用が有効だとされており (McNairn *et al.*, 2009)、既往の研究でも PALSAR-2 などの Lバンドによる作付作物分類が可能であることが示されている (Lee *et al.*, 2001; Whelen and Siqueira, 2018)。また、この Lバンドを Cバンドと併用した研究例も報告されている。1 時期での土地被覆分類においては Xバンドを使用しなくてもこの 2バンドのみで高い精度を得られることも報告されており (Qi *et al.*, 2019)、Lバンドが他の波長の SAR の代替として利用可能であることを示唆している。しかし、畑作地帯の作付作物分類において、多時期の RADARSAT-2 と PALSAR-2 を併用することの有効性が示されている研究はみられず、検討が必要であると考えられる。また、各種の散乱成分分解手法から算出される散乱成分を併用することにより、作付作物分類の精度が向上することを

示した研究がなされており (Xu *et al.*, 2014), 2 章や 3 章でも散乱成分の有効性を示した。RADARSAT-2 や PALSAR-2 を使用した分類において、後方散乱係数だけでなく散乱成分を算出して分類に使用した例もみられない。すなわち、後方散乱係数と散乱成分を使用した分類において、アーカイブで取得できる L バンド SAR データと C バンド SAR データの各分類結果の比較を行い、多時期の分類において両者を併用することの有効性を示す必要があるといえる。既往の研究では、1 時期の SAR のみを使用した土地被覆分類は行われているが (Qi *et al.*, 2019), 多時期の SAR データを使用した作付作物分類において、使用する衛星データ数の削減を検討した研究はあまりみられない。PALSAR-2 が RADARSAT-2 よりもデータの取得コストが低いことを重要視し、L バンド SAR データと一部の時期の C バンド SAR データを併用した分類を行い、低コストかつ高精度な分類を目指すためのデータ選択方法を提示する必要があるといえる。

本章では、C バンド SAR である RADARSAT-2 の 4 偏波データと、L バンド SAR である ALOS2/PALSAR-2 の単偏波データおよび 4 偏波データを使用し、畑作地帯における作付作物の分類を行った。PALSAR-2 データについては、アーカイブからデータを選択し使用した。分類を行うにあたり、先述のとおり SAR データから算出される主要な数値の後方散乱係数を使用した。RADARSAT-2 や PALSAR-2 の 4 偏波データからは、3 章で最も高い精度の組み合わせとして得られた後方散乱係数、van Zyl (2011) による散乱モデル分解法の散乱成分、Cloude and Pottier (1996) による固有値解析法の散乱成分の 3 データセットを算出し、作付作物の分類処理の変数として使用した。PALSAR-2 の単偏波データからは、HH 偏波の後方散乱係数を算出して使用した。分類の手法には、2, 3 章と同様にパターン認識の手法を用いて高い精度が期待できる機械学習アルゴリズムの 1 つである Random Forests (RF) (Breiman, 2001) を使用した。

以上の点を踏まえ、まず RADARSAT-2 データや PALSAR-2 データから得られた後方散乱係数や散乱成分の 9 変数を用いて、様々な条件下でそれぞれの衛星データを用いた RF による分類を行い、分類に有効な SAR データの波長や時期、変数の検討から、高精度な分類手法を探ることを第一の目的とした。次に、両波長のデータを併用した分類を行い、9 変数を用いて 2 つの波長を併用することの有効性について検討した。また先述のとおり、PALSAR-2 は RADARSAT-2 より取得コストが低いですが、実利用の観点からはできるだけ少ない SAR データで高い精度を得る手法を検討する必要があるため、本章では SAR の時期数を削減した

分類や、RADARSAT-2 を PALSAR-2 に置換した分類を実施することにより、できるだけコストを削減した効率的な分類手法を検討した。

## 5.2 研究方法

### 5.2.1 使用データおよび処理方法

本章の対象地域および対象作物は2章～4章と同じである。使用した衛星データは、Cバンドのセンサを搭載した RADARSAT-2 および L バンドのセンサを搭載した ALOS-2/PALSAR-2 のデータである。RADARSAT-2 は3章と同様に4偏波データを使用し、PALSAR-2 は単偏波 (HH 偏波) データ (観測モード: UBS) と4偏波データ (観測モード: HBQ) を使用した。

データを取得した時期は2015年と2017年の2か年である。2015年は、5時期のRADARSAT-2 と2時期の4偏波PALSAR-2 を取得した。2017年は、6時期のRADARSAT-2 と2時期の単偏波PALSAR-2, 1時期の4偏波PALSAR-2 を取得した。表5-1, 5-2に、それぞれの衛星データの観測日とRADARSAT-2の入射角、PALSAR-2のオフナディア角を示す。以降、各年の観測時期を早い方から1期～5期, 1期～6期と記述する。

RADARSAT-2の周期は24日間隔であるため、2015年は5月下旬から8月下旬までの同じ入射角で観測されたデータをすべて取得した。2017年は、使用するデータの観測時期の間隔を狭めるため、6月上旬から8月下旬までに観測された異なる2種類の入射角のデータ

表 5-1 2015 年の SAR データ 取得時期

時期	RADARSAT-2		PALSAR-2		
	観測日	入射角	観測日	偏波	オフナディア角
1期	5月26日	39.2° ~40.7°	8月10日	4偏波	25.5°
2期	6月19日	39.2° ~40.7°	9月26日	4偏波	33.2°
3期	7月13日	39.2° ~40.7°			
4期	8月 6日	39.2° ~40.7°			
5期	8月30日	39.2° ~40.7°			

表 5-2 2017 年の SAR データ 取得時期

時期	RADARSAT-2		PALSAR-2		
	観測日	入射角	観測日	偏波	オフナディア角
1期	6月 8日	39.2° ~40.7°	6月 3日	単偏波	32.8°
2期	6月22日	29.1° ~30.9°	6月24日	単偏波	35.8°
3期	7月 2日	39.2° ~40.7°	8月 7日	4偏波	25.4°
4期	7月16日	29.1° ~30.9°			
5期	7月26日	39.2° ~40.7°			
6期	8月 9日	29.1° ~30.9°			

を取得した。ピクセルスペーシングや処理レベルは 3 章と同様である。

PALSAR-2 は、2015 年、2017 年ともに、RADARSAT-2 の観測時期の範囲のデータをアーカイブからすべて取得した。アーカイブに含まれているデータは、一般利用する際であってもいつでも購入できるものである。ピクセルスペーシングは単偏波が 3m、4 偏波が 6m で、2017 年の 6 月 24 日のみディセンディング軌道、他は全てアセンディング軌道より観測されたものである。処理レベルはすべて Single Look Complex (SLC) である。

衛星観測のタイミングをこのように決定し、取得されたデータを用いて、分類に使用する変数を算出した。RADARSAT-2 については、4 章と同じ 3 手法より算出される後方散乱係数および散乱成分の 3 種類のデータセットを使用した (表 5-3)。この 3 手法の 9 要素を 2015 年は 5 時期、2017 年は 6 時期のデータから算出した。PALSAR-2 の 4 偏波データ (2015 年 8 月 10 日、9 月 26 日、2017 年 8 月 7 日) についても同様に、これらの 3 手法より算出される 9 要素を算出して使用した。単偏波データ (2017 年 6 月 3 日および 6 月 24 日) については、散乱成分を算出することができないため、HH 偏波の後方散乱係数の 1 変数のみ算出し

表 5-3 算出した各データセットの要素

衛星データ	データセット	算出した要素
4偏波RADARSAT-2 4偏波PALSAR-2	$\sigma^0$	HH偏波 ( $\sigma_{HH}^0$ ) HV偏波 ( $\sigma_{HV}^0$ ) VV偏波 ( $\sigma_{VV}^0$ )
	CP	エントロピ (H) アニソトロピ (A) アルファ角 ( $\alpha$ )
	VZ	二回散乱 (Pd) 体積散乱 (Pv) 表面散乱 (Ps)
単偏波PALSAR-2	$\sigma^0$	HH偏波 ( $\sigma_{HH}^0$ )

て使用した。各衛星データについて、前章までと同様の方法でスペックルフィルタやマルチルック処理、オルソ補正を行った。

## 5.2.2 解析方法

本章でも 2 章と同様に、対象地域について十勝農業共済組合より提供を受けた圃場情報 GIS データを用いて、バッファ処理や平均値の抽出を行った。分類処理も同様に、ホールドアウト法により圃場を 20%の教師データと 80%のテストデータに分けて実行し、このテストデータに対する分類結果を基に精度を評価した。教師データとテストデータに用いた作物種別の圃場数を表 5-4 に示す。

分類手法も 2 章と同様に Random Forests (RF) を使用した。使用する決定木の本数は 4 章と同様に 500 とした。本章でも、分類の精度の評価には全体精度 (OA)、カッパ係数 ( $\kappa$ )、プロデューサーズ精度 (PA)、ユーザーズ精度 (UA) の 4 種類を使用した。2 つの手法間における  $\kappa$  の差の有意性については、検定統計量 Z を使用して評価した。

## 5.2.3 解析の流れ

本章では、以下の 1) ~ 3) で示す 3 処理で解析を進めた。

### 1) 1 時期のデータを使用した分類でのバンド間の比較

1 時期の RADARSAT-2 データを使用した分類、PALSAR-2 データを使用した分類、両者を併用した分類のそれぞれにおいて、後方散乱係数と散乱成分の 9 変数を用いた分類にお

表 5-4 各圃場数の内訳

作物	2015年			2017年		
	教師データ	テストデータ	合計	教師データ	テストデータ	合計
小麦	272	1,086	1,358	275	1,102	1,377
豆類	195	779	974	189	756	945
馬鈴薯	143	573	716	133	530	663
甜菜	107	426	533	112	448	560
牧草	120	478	598	119	475	594
トウモロコシ	122	490	612	115	459	574
合計	959	3,832	4,791	943	3,770	4,713

ける精度を確認しておく必要がある。そこで、RADARSAT-2 データと PALSAR-2 データのうち、条件や観測時期の近いデータを双方で 1 時期ずつ選定し、各々を使用した分類結果を比較した。ここでは、4 偏波であることが共通する 2015 年に観測された RADARSAT-2 の 4 期（8 月 6 日）と PALSAR-2 の 1 期（8 月 10 日）を使用した。したがって、異なるバンドの衛星データで同種のデータセット（後方散乱係数、散乱モデル分解法、固有値解析法による 9 変数）を使用した場合の精度の差異についての検討ができる。

また、この RADARSAT-2 の 4 期と PALSAR-2 の 1 期を併用することにより、片方の波長のデータのみを使用した分類と比較して精度の向上がみられるかについても併せて検討した。

## 2) 多時期のデータを使用した分類でのバンド間の比較

3 章によると、多時期のデータを使用し、後方散乱係数と散乱成分の 9 変数を用いた分類は、精度の向上をもたらした。そこで、両バンドで観測されたデータのうち複数の時期のものを使用し、分類結果の比較を行った。まず、2015 年のデータについて、それぞれのバンドですべての時期のデータを使用した。RADARSAT-2 については 6 時期のデータによる分類、PALSAR-2 については 2 時期の 4 偏波データによる分類をそれぞれ行い、バンド間で精度を比較した。

また、1) と同様に、この RADARSAT-2 の 6 時期のデータと PALSAR-2 の 2 時期のデータを併用することによる精度の変化を検討した。

次に、この 2015 年の結果を踏まえて、2017 年のデータでも分類結果の比較を行った。2015 年のデータは、両バンドで取得した時期や時期数が大幅に異なっていたが、2017 年は PALSAR-2 の単偏波のデータを取得したことで、両者の観測時期や時期数の条件を揃えやすくなり、より多時期での分類における精度の比較が可能となる。RADARSAT-2 データは、1, 2, 6 期（6 月 8 日、6 月 22 日、8 月 9 日）を使用して分類を行い、PALSAR-2 データは、1, 2, 3 期（6 月 3 日、6 月 22 日、8 月 7 日）すべてを使用して分類を行った。これらの精度の比較のほか、両者を併用した分類についても比較を行った。

## 3) 一部の RADARSAT-2 を PALSAR-2 で代替した分類

RADARSAT-2 は PALSAR-2 に比べて取得に費用がかかるため、実利用の観点から RADARSAT-2 をできるだけ使用せずに PALSAR-2 で代替することの可能性に注目して検討した。まず、6 時期の RADARSAT-2 データのうちいくつかを PALSAR-2 データに変更する

手法により、使用するシーン数を6時期に固定した分類で精度の比較を行った。すなわち、RADARSAT-2データの1, 2, 6期のうちのいずれかを、それぞれ観測時期に近いPALSAR-2データの1, 2, 3期に変更し、精度の差異を比較検討した。

次に、2~3時期のPALSAR-2データに、1~2時期のRADARSAT-2データを追加する手法により、使用するシーン数を4時期に固定した分類で精度の比較を行った。1点目のパターンとして、3時期すべてのPALSAR-2データに、1時期のRADARSAT-2データを加えて分類を実行した。なおRADARSAT-2データは、精度の向上のため、PALSAR-2による取得が行われなかった7月に撮影された3時期(3, 4, 5期)のうちいずれか1時期を加えることとした(計3種類)。2点目のパターンとして、2時期のPALSAR-2データに、2時期のRADARSAT-2データを加えて分類を実行した。3章において、6月から9月までのRADARSAT-2データのうち1時期を使用した作付作物分類を行い、時期ごとの精度を比較した結果、Cバンドのデータは早い時期の方が高い精度を得られることが示されている。そこで、PALSAR-2データは2, 3期のみを使用し、RADARSAT-2データは1期と7月(3, 4, 5期)のうちいずれか1時期の計2時期を使用することとした(計3種類)。

## 5.3 結果と考察

### 5.3.1 1時期を使用した分類でのバンド間の比較

2015年におけるRADARSAT-2の4期(8月6日)とPALSAR-2の1期(8月10日)のデータを使用した場合における各々の分類結果を比較したものを表5-5に示す。RADARSAT-2に関しては、1時期のみで0.6を上回るOAで分類がなされていた。とりわけ豆類、甜菜、小麦はPA, UAともに0.7を上回るOAが達成された。一方、PALSAR-2はOA, κともにRADARSAT-2による分類を大幅に下回る結果となっており、特に豆類や甜菜はPA, UAとも極めて低い精度となった。

バンドにより精度に差異が生じた原因として、各バンドで作物体を透過する度合いが異なる点が考えられる。CバンドであるRADARSAT-2は、LバンドであるPALSAR-2と比較すると作物体を透過しにくく、豆類や甜菜のように条間は広くても葉が多く水平に広がる作物を捉えやすい可能性が考えられる。また、馬鈴薯は黄変、牧草は収穫により、この時期

表 5-5 2015 年の 1 時期のデータを使用した分類の精度

	RADARSAT-2 8月6日		PALSAR-2 8月10日		併用	
	PA	UA	PA	UA	PA	UA
豆類	0.770	0.729	0.302	0.288	0.791	0.724
甜菜	0.721	0.785	0.178	0.287	0.737	0.803
牧草	0.674	0.714	0.385	0.426	0.718	0.757
トウモロコシ	0.486	0.517	0.367	0.413	0.592	0.655
馬鈴薯	0.485	0.506	0.318	0.332	0.503	0.589
小麦	0.806	0.756	0.592	0.482	0.860	0.775
OA	0.684		0.391		0.727	
$\kappa$	0.609		0.240		0.661	

は圃場ごとに状態のばらつきがあるため、1 時期のみでの分類は比較的難しかった可能性が考えられる。

一方、L バンドは作物体を透過しやすいため、豆類や甜菜のような作物は透過されて作物の形状を捉えにくく、かつ条間も広いために土壌の影響を受けやすいことで、判別が困難であったと考えられる。牧草は作物体同士の間隔が狭く密に植えられているため、マイクロ波の散乱によってそれ以外の作物との差異をある程度判別できたと考えられる。

表 5-5 の右列には、RADARSAT-2 の 4 期 (8 月 6 日) と PALSAR-2 の 1 期 (8 月 10 日) を併用した分類結果も示した。同表によると、併用することにより RADARSAT-2 のみでの分類結果と比較して OA、 $\kappa$  ともに向上しており、さらに両者間で  $\kappa$  に有意な差がみられた。

RADARSAT-2 のみでは判別がしにくかった牧草と馬鈴薯、トウモロコシの UA や PA が向上し、精度が元から高かった小麦についても PA の向上がみられた。この時期は一部の牧草と多くの小麦で収穫がなされているが、小麦は登熟期での収穫、牧草は出穂直後での収穫と、両作物の水分量などの状態が異なる。C バンドと L バンドを併用することで、両者の水分量の差異に各バンドの散乱が影響されて小麦や牧草の PA、UA 精度が向上したと考えられる。馬鈴薯は畝がありかつ黄変が進んでいる時期であり、トウモロコシは生育が他の作物より進んでいるために、バンド間で透過の度合いが異なるとそれぞれのバンドで他の作物と散乱の違いが明瞭となり、精度の上昇に寄与したと考えられる。以上のように、1 時期での

作付作物分類では、RADARSAT-2 データと PALSAR-2 データを併用することにより、一方の波長のみを使用した分類と比較して有意差のある  $\kappa$  が得られることが期待できる。

### 5.3.2 多時期を使用した分類でのバンド間の比較

2015 年における RADARSAT-2 の全 5 時期を使用した分類結果と、PALSAR-2 の全 2 時期を使用した分類結果を各々比較したものを表 5-6 に示す。RADARSAT-2 を 5 時期使用した場合、0.9 を超える OA の分類が可能となった。PALSAR-2 による分類でも、2 時期のデータを使用することで、5.3.1 で示した 1 時期のみを使用した分類と比較すると精度の向上がみられ、 $\kappa$  にも有意差がみられたが、2 時期の PALSAR-2 を使用しても、OA、 $\kappa$  ともに 5.3.1 で示した 1 時期の RADARSAT-2 による分類より低く、 $\kappa$  にも有意な差がみられた。5.3.1 の結果と同様に、波長の短い C バンドの RADARSAT-2 の方が L バンドの PALSAR-2 と比較して作物体の透過が少ないために、作物種ごとの判別がしやすかったためであると考えられる。4 章の結果では、7 月上旬までの RADARSAT-2 と、7 月中旬以降の X バンドである TerraSAR-X のデータを併用することで、最も高い精度での分類がなされた。すなわち、時期が後半になると短い波長の衛星で作物を判別しやすくなる可能性が示唆された。したがって、L バンドである PALSAR-2 のような長波長のデータは、より早い時期のデータを使用することで精度の向上に寄与できる可能性が考えられる。

表 5-6 2015 年の多時期のデータを使用した分類の精度

	RADARSAT-2 5時期		PALSAR-2 2時期		併用	
	PA	UA	PA	UA	PA	UA
豆類	0.883	0.915	0.336	0.356	0.896	0.915
甜菜	0.892	0.929	0.477	0.537	0.897	0.918
牧草	0.939	0.914	0.464	0.561	0.937	0.916
トウモロコシ	0.822	0.796	0.439	0.467	0.831	0.829
馬鈴薯	0.901	0.918	0.389	0.507	0.911	0.927
小麦	0.984	0.961	0.667	0.509	0.983	0.961
OA	0.915		0.483		0.920	
$\kappa$	0.895		0.353		0.901	

作物別では、PALSAR-2において、甜菜は1時期での分類に比べPA・UAとも精度の改善がみられた。この理由として、甜菜は地上部の作物体が大きい9月下旬のデータが加わったことで、他の作物と散乱の違いが明瞭となり、判別がしやすかったためであると考えられる。また、牧草の誤分類も大幅に減少した。この時期の牧草は生長しているが、小麦や馬鈴薯、一部のトウモロコシや豆類は既に収穫が終了しているため、他の作物への誤分類が減少したことで精度が向上したと考えられる。一方、全てあるいは一部の圃場で既に収穫が終了しているこれらの4作物ではPAやUAが向上しにくく、高いOAが得られなかった原因であると考えられる。

表5-6の右列には、2015年の全5時期のRADARSAT-2と全2時期のPALSAR-2を併用した分類結果も示した。同表によると、RADARSAT-2のみを使用した場合より高い精度が得られたが、両者の $\kappa$ に有意な差はみられなかった。また、ジニ係数減少量を用いて変数の重要度を確認したところ(図5-1)、上位40変数はすべてRADARSAT-2の変数であったことからも、今回の条件ではPALSAR-2を併用しても精度向上への効果は低いと考えられる。

つづいて、多時期のデータを使用した分類の比較を行うために、2017年のRADARSAT-2とPALSAR-2を用いてそれぞれ分類を行った。まず、各バンドの精度を比較するため、双方のバンドで観測時期の近いデータを3時期選定し、各々の精度を算定した。RADARSAT-2データは、1, 2, 6期(6月8日, 6月22日, 8月9日)を使用し、PALSAR-2データは、1, 2, 3期(6月3日, 6月22日, 8月7日)すべてを使用した。結果の判別効率表を表5-7, 5-8に示す。

RADARSAT-2データでの分類では、0.9に近いOAで分類がなされており、XバンドのTerraSAR-Xを3時期使用した2章の結果より高い精度が得られた。とくにトウモロコシ以外の作物はPA, UAともに0.8を上回る精度が達成された。トウモロコシに関しては、RADARSAT-2のみでの判別が難しいことが示され、3章と同様の結果となった。一方、PALSAR-2データでの分類では、OAが0.573となり、RADARSAT-2より大幅に低い精度となった。

作物別に2017年のPALSAR-2データによる分類を見ると、甜菜はPA, UAとも極めて低い精度であった。一方、豆類は2015年のPALSAR-2データでの分類において低いPA, UAとなっていたが、2017年の分類では大幅な精度の向上がみられた。これは、豆類の作物体がまだ小さい6月のデータを使用したことにより、既に生育が進んでいる甜菜や馬鈴薯

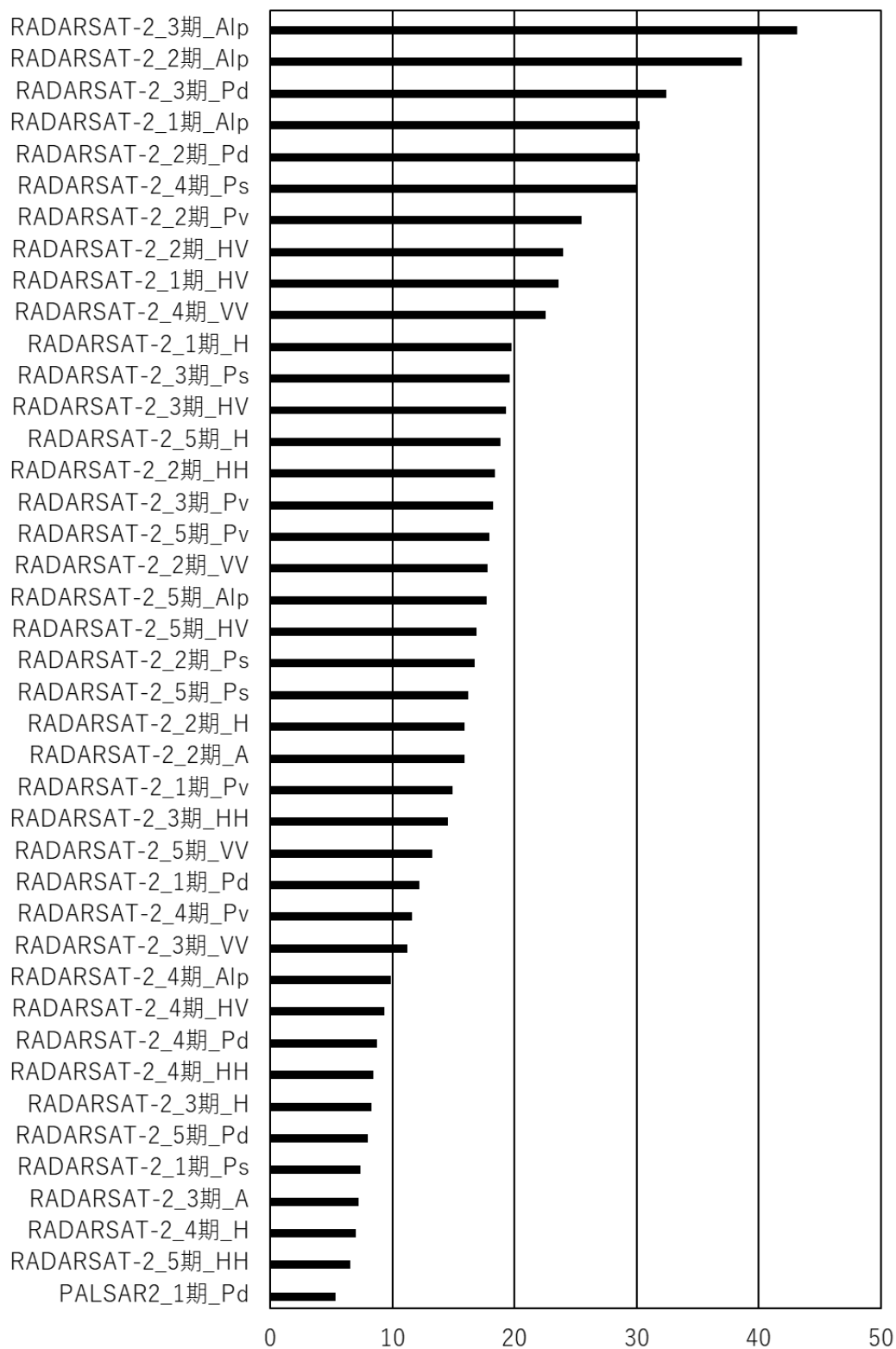


図 5-1 2015 年の PALSAR-2 および RADARSAT-2 データを 1 時期ずつ併用した分類におけるジニ係数減少量

表 5-7 2017 年の RADARSAT-2 データ 1, 2, 6 期目を使用した分類による判別効率表

		参照データ							
		豆類	甜菜	牧草	トウモロコシ	馬鈴薯	小麦	合計	UA
分類結果	豆類	681	6	7	73	3	10	780	0.873
	甜菜	12	401	6	10	16	4	449	0.893
	牧草	2	0	412	19	21	8	462	0.892
	トウモロコシ	47	5	16	321	14	9	412	0.779
	馬鈴薯	9	33	6	27	473	5	553	0.855
	小麦	5	3	28	9	3	1,066	1,114	0.957
	合計	756	448	475	459	530	1,102	3,770	
	PA	0.901	0.895	0.867	0.699	0.892	0.967		
								全体精度	0.890
								カッパ係数	0.864

表 5-8 2017 年の PALSAR-2 データ 1, 2, 3 期目を使用した分類による判別効率表

		参照データ							
		豆類	甜菜	牧草	トウモロコシ	馬鈴薯	小麦	合計	UA
分類結果	豆類	493	54	62	94	62	85	850	0.580
	甜菜	14	143	13	11	57	46	284	0.504
	牧草	26	19	210	35	10	66	366	0.574
	トウモロコシ	57	14	36	215	21	38	381	0.564
	馬鈴薯	48	71	16	32	290	56	513	0.565
	小麦	118	147	138	72	90	811	1,376	0.589
	合計	756	448	475	459	530	1,102	3,770	
	PA	0.652	0.319	0.442	0.468	0.547	0.736		
								全体精度	0.573
								カッパ係数	0.464

などとの判別がしやすくなったためであると考えられる。

以上の結果でも 2015 年の結果と同様に、C バンドである RADARSAT-2 は PALSAR-2 と比較すると作物体を透過しにくいため、作物体の違いを捉えられやすかったと考えられる。L バンドである PALSAR-2 は、作物体を透過しやすいため、RADARSAT-2 と比較すると豆

類や甜菜のような作物の形状を捉えにくかったと考えられる。一方で小麦や牧草は作物体同士の間隔が狭く密に植えられているため、マイクロ波の散乱によってそれ以外の作物との差異をある程度判別できたと考えられる。また馬鈴薯は、作物体を透過しても、畝方向がレーダ照射方向と垂直に近い場合に散乱が大きくなるため、圃場の向きによっては他の作物との判別がされやすく、比較的高い PA や UA が得られたと考えられる。

2015 年の結果と 2017 年の結果を比較すると (表 5-6, 5-8), 2015 年の PALSAR-2 データを使用した分類の OA は 0.48 であったが、2017 年の結果は 8 月以前の時期のデータを使用したことにより、精度の向上が示された。分類における変数の重要度 (ジニ係数減少量) を算出したところ、1 期と 2 期の HH 偏波が特に重要度の高い 2 変数であった。したがって、これらの時期の情報が作付作物の分類に大きく寄与したことが考えられる。しかし、3 章で示したように、HV 偏波や各種の散乱成分は作付作物分類の精度向上に寄与することから、単偏波ではなく 4 偏波での観測データを使用することによって、さらに精度の向上が見込まれると考えられる。

RADARSAT-2 データの 1, 2, 6 期と PALSAR-2 データの 1, 2, 3 期を併用した分類結果を表 5-9 に示し、その結果を元に作成した作付マップを図 5-2 に示す。併用により、

表 5-9 2017 年の RADARSAT-2 データおよび PALSAR-2 データを併用した分類による判別効率表

		参照データ							
		豆類	甜菜	牧草	トウモロコシ	馬鈴薯	小麦	合計	UA
分類結果	豆類	686	6	6	73	4	11	786	0.873
	甜菜	10	404	5	10	12	5	446	0.906
	牧草	2	1	405	20	20	7	455	0.890
	トウモロコシ	43	3	18	329	12	5	410	0.802
	馬鈴薯	10	31	9	20	479	4	553	0.866
	小麦	5	3	32	7	3	1,070	1,120	0.955
	合計	756	448	475	459	530	1,102	3,770	
PA		0.907	0.902	0.853	0.717	0.904	0.971		
		全体精度							0.895
		カッパ係数							0.870

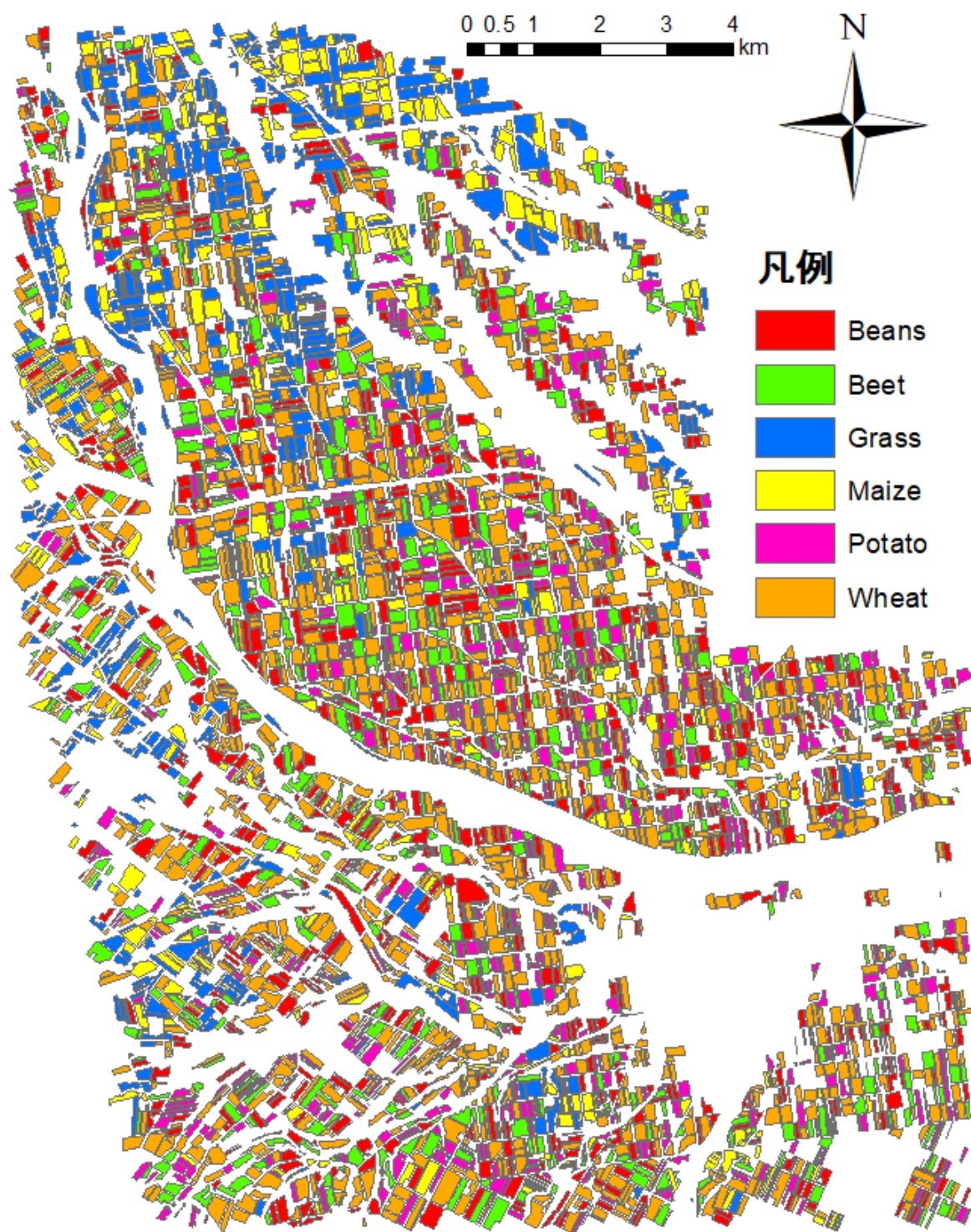


図 5-2 2017 年 RADARSAT-2 データと PALSAR-2 データを併用した分類による  
作付マップ

RADARSAT-2 データのみでの分類結果 (表 5-7) と比較して OA,  $\kappa$  ともに向上したが, 両者間で  $\kappa$  に有意な差はみられなかった。ジニ係数減少量を用いて変数の重要度を確認したところ (図 5-3), 1 期の PALSAR-2 の HH 偏波が 21 番目に高い値であったが, 上位 20 変数はすべて RADARSAT-2 の変数であったことから, 6 月の単偏波 PALSAR-2 データを加えても, 2015 年の結果と同様に併用する必要性は低いと考えられる。

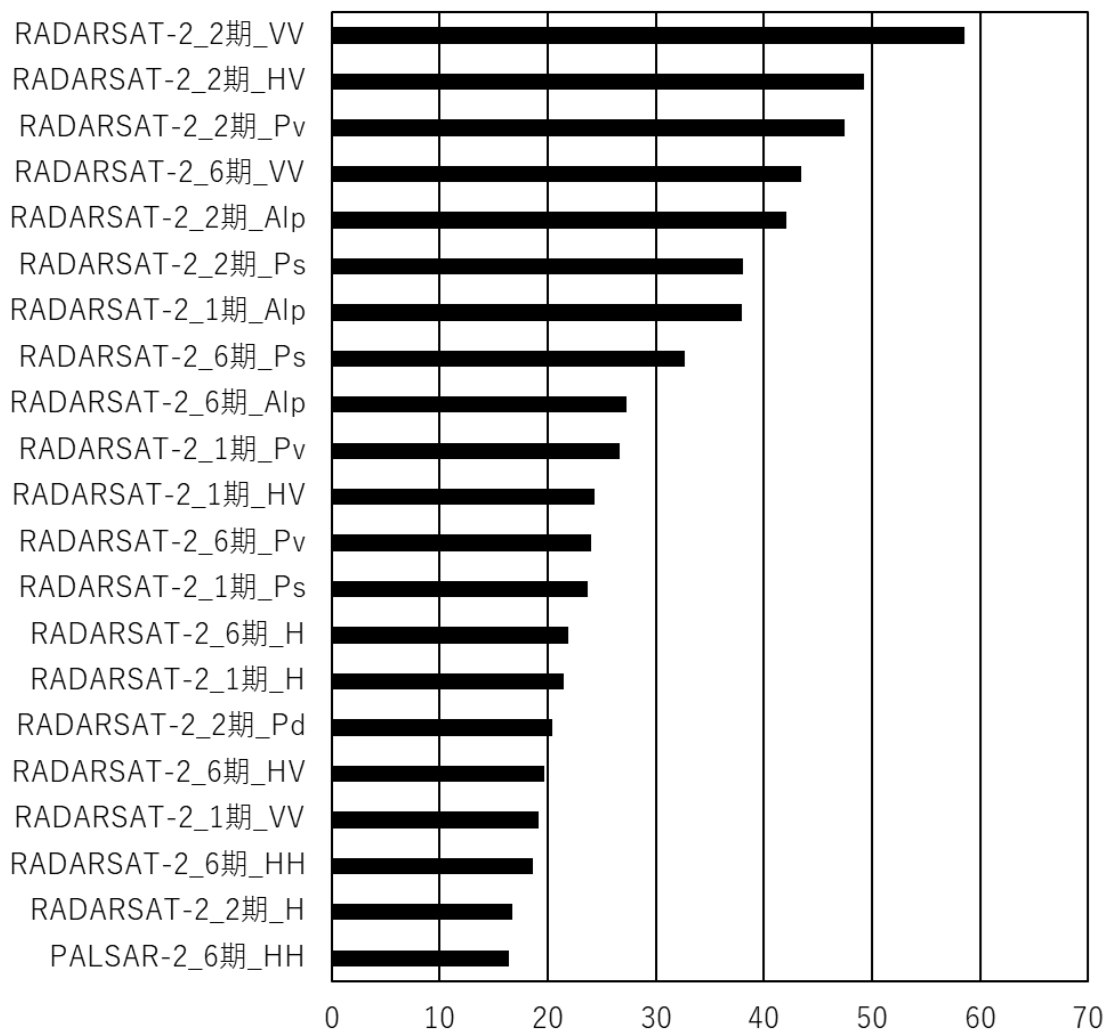


図 5-3 2017 年の PALSAR-2 および RADARSAT-2 データを 3 時期ずつ併用した分類におけるジニ係数減少量

### 5.3.3 一部の RADARSAT-2 を PALSAR-2 で代替した分類

RADARSAT-2 は PALSAR-2 に比べてデータ取得に費用がかかるため、実利用の観点から RADARSAT-2 をできるだけ使用せずに PALSAR-2 で代替することの実用可能性に注目して検討した。まず、6 時期の RADARSAT-2 データのうちいくつかを PALSAR-2 データに置換する手法により、使用するシーン数を 6 時期に固定した分類により精度の比較を行った。RADARSAT-2 データの 1, 2, 6 期のうちのいずれかを、それぞれ観測時期の近い PALSAR-2 データの 1, 2, 3 期に変更し、精度の差異を比較検討した結果が表 5-10 である。以下では、表の最下行に示した RADARSAT-2 データを 6 時期すべてを使用した分類結果との比較を中心に考察する。

RADARSAT-2 データの 5 時期に PALSAR-2 データの 1 時期を併用した分類は、若干の精度低下はみられたものの、OA はすべて 0.9 を超えており、それぞれの  $\kappa$  に有意な差はなかった。この中では RADARSAT-2 データの 1 期を PALSAR-2 データの 1 期に変更した分類が最も高い精度となり、RADARSAT-2 データのみを使用した分類とほぼ同じ OA や  $\kappa$  となった。

次に、RADARSAT-2 データの 4 時期に PALSAR-2 データの 2 時期を併用した分類では、精度がさらに低下し、RADARSAT-2 を 6 時期使用した分類と比較して  $\kappa$  に有意な差がみられた (表 5-10)。しかし、RADARSAT-2 の 1, 6 期を PALSAR-2 に変更した分類 (表 5-10 の

表 5-10 2017 年の RADARSAT-2 データの一部を PALSAR-2 データに変更した  
計 6 時期による分類

使用した RADARSAT-2(期)	使用した PALSAR-2(期)	OA	$\kappa$
2, 3, 4, 5, 6	1	0.911	0.890
1, 3, 4, 5, 6	2	0.910	0.888
1, 2, 3, 4, 5	3	0.909	0.888
3, 4, 5, 6	1, 2	0.888	0.861
2, 3, 4, 5	1, 3	0.902	0.879
1, 3, 4, 5	2, 3	0.905	0.882
3, 4, 5	1, 2, 3	0.882	0.854
1, 2, 3, 4, 5, 6	なし	0.919	0.900

5行目)と、RADARSAT-2の2, 6期をPALSAR-2に変更した分類(表5-10の6行目)は、OAが0.9を超えた。これらは、6月の1時期と8月の1時期をPALSAR-2に変更したものである。RADARSAT-2の1, 2期をPALSAR-2に変更した分類は、6月の2時期を変更しているが、0.9を超えるOAが得られなかった。このことから、6月は単偏波のPALSAR-2だけでは精度が低下するため、RADARSAT-2を1時期は確保することが必要であるといえる。この原因として、単偏波の衛星データでは算出できる変数が少なく、分類を行うために十分な数の変数を確保できなかったことが考えられる。8月のPALSAR-2は4偏波のため、HVや散乱成分などの重要な変数を算出でき、比較的高いOAが得られたと考えられる。RADARSAT-2データの3時期にPALSAR-2データの3時期を併用した分類のOAは0.882とさらに低下した(表5-10)。

つづいて、2~3時期のPALSAR-2に1~2時期のRADARSAT-2を加え、合計4時期を使用して分類した結果を表5-11に示す。PALSAR-2データを3時期すべてを使用した分類(表5-8)と、PALSAR-2の3時期にRADARSAT-2の7月に取得された1時期を併用した分類とを比較したところ、PALSAR-2のみ3時期の分類では、0.573と高いOAは得られなかったが、これらに7月のRADARSAT-2を1時期加えるだけで大幅に精度の向上(0.770~0.824)がみられた。とくに、RADARSAT-2の3期である7月2日のデータを加えることによって、OAは0.8を超える結果が得られた。4章では、7月上旬までのRADARSAT-2データと、7月中旬以降のXバンドであるTerraSAR-Xデータを併用することで、最も高い精度での分類

表 5-11 2017 年における RADARSAT-2 データ 1~2 時期と  
PALSAR-2 データ 2~3 時期を併用した計 4 時期による分類

使用した RADARSAT-2(期)	使用した PALSAR-2(期)	OA	$\kappa$
3	1, 2, 3	0.824	0.783
4	1, 2, 3	0.780	0.726
5	1, 2, 3	0.770	0.713
1, 3	2, 3	0.877	0.848
1, 4	2, 3	0.852	0.817
1, 5	2, 3	0.843	0.805

がなされ、早い時期の C バンドデータを用いることで作物を判別しやすくなる可能性が示唆されていた。本章でも、RADARSAT-2 のデータは、より早い時期のものを使用することが精度の向上に寄与すると考えられる。

この結果を踏まえ、6 月上旬の早い時期の RADARSAT-2 データが必要であると判断されるため、PALSAR-2 の 1~3 期のうち 1 期を RADARSAT-2 に変更し、PALSAR-2 データを 2, 3 期の 2 時期と RADARSAT-2 データを 2 時期使用した分類も実行した (表 5-11)。その結果、更なる精度の向上がみられた。とくに RADARSAT-2 データの 1, 3 期と PALSAR-2 データの 2, 3 期を使用することで、0.877 の OA が得られた。こちらの結果でも、より早い時期のデータを使用することで精度が高くなったといえる。

#### 5.4 まとめ

本章では、C バンドの 4 偏波 SAR である RADARSAT-2 データと、L バンドの単偏波 SAR および 4 偏波 SAR である PALSAR-2 データを併用し、後方散乱係数や散乱成分の計 9 変数を用いた作付作物の分類を実行した。

まず、8 月上旬に観測された 1 時期の 4 偏波データで両者を比較したところ、RADARSAT-2 データの方が高い OA を得られた。PALSAR-2 データからは、良好な OA を得られず、両者の  $\kappa$  に有意差がみられた。しかし、両者を併用することにより OA は向上し、RADARSAT-2 データのみを使用した分類と比較して  $\kappa$  に有意差がみられた。以上より、1 時期の後方散乱係数や散乱成分の 9 変数を使用した作付作物分類では、PALSAR-2 データよりも RADARSAT-2 データの方が良好な精度が得られるが、RADARSAT-2 と PALSAR-2 のデータを併用することにより、各バンドが地表面の状態に対して異なる散乱をするために、情報を補完し合うことで精度の向上が期待できることから、併用することの有効性が示された。

次に、各バンドで観測されたデータのうち複数の時期のものを使用し、分類結果の比較を行った。2015 年の結果については、RADARSAT-2 では 0.9 を超える OA を得られたのに対し、PALSAR-2 では 0.5 を下回る結果となった。また、両者を併用した分類でも、RADARSAT-2 のみを使用した分類と同等の精度であった。ただし、2015 年の条件下では、PALSAR-2 データを 8 月以降の 2 時期しか使用していない。すなわち、両者を比較するにあたり、PALSAR-

2 のデータが不足していたことも高い精度が得られなかった原因であると考えられる。そこで 2017 年については、6 月に観測された単偏波の PALSAR-2 データ 2 時期を取得し、8 月の 4 偏波データと併せて計 3 時期のデータで分類を行い、精度の向上や効率化を検討した。また、RADARSAT-2 データも、PALSAR-2 で観測されたデータの時期に近い 3 時期のみを使用し、条件を近づけて比較した。その結果、RADARSAT-2 データによる分類の方が精度は高かったが、PALSAR-2 データについても、2015 年の結果より大幅な精度の向上がみられた。しかし、両者を併用した分類では RADARSAT-2 のみを使用した分類と  $\kappa$  に有意差はみられなかった。以上のことから、多時期による分類では、PALSAR-2 の少ない変数による分類よりも RADARSAT-2 による 9 変数での分類の方が高い精度となり、PALSAR-2 による分類が有効でないことを示した。また、RADARSAT-2 で 9 変数を算出することで、PALSAR-2 データは併用する必要性が低いことも示した。

さらに、RADARSAT-2 をできるだけ使用せずに PALSAR-2 で代替することが可能であるかに注目して検討した。まず、6 時期の RADARSAT-2 データのうち、いくつかを PALSAR-2 データに変更する手法を実行した。その結果、1 時期を PALSAR-2 データに変更しても  $\kappa$  に有意差のない精度が得られた。2 時期を変更すると  $\kappa$  に有意差がみられたが、6 月の 1 時期と 8 月の 1 時期を PALSAR-2 データに変更した分類では 0.9 を超える OA が得られた。次に、2~3 時期の PALSAR-2 データに 1~2 時期の RADARSAT-2 データを加える手法を実行したところ、PALSAR-2 データのみでの分類と比較して大幅に精度の向上がみられた。とくに 7 月上旬までの RADARSAT-2 データを 1~2 時期加えて分類することにより、効率的に高い精度を得ることができた。

以上のことから、2 時期の PALSAR-2 データと、6 月のデータを含めた 4 時期の RADARSAT-2 データを併用することにより、0.9 を超える精度で分類を行うことができ、2 章で示した X バンド SAR による分類よりも高精度であったことから、取得コストを削減した効率的な分類を達成できると結論づけた。また、PALSAR-2 を中心とした分類は、2 時期の PALSAR-2 データに 6 月や 7 月上旬の RADARSAT-2 データを 2 時期追加する手法は、少ないシーン数でも高い精度を得られることから、実用的に L バンド SAR データを分類に使用する方法となりうることを示された。

C バンドや L バンドのような長波長のデータは、より早い時期のものを使用することで精度の向上に寄与すると考えられるが、本章で使用した PALSAR-2 データのアーカイブで

は 6 月など早い時期のデータが単偏波のものだけであった。しかし、本章での結果から、HV 偏波やエントロピなどは作付作物を判別するために重要な変数であるため、PALSAR-2 の 4 偏波データで 9 変数を算出し、RADARSAT-2 データと併用することで、分類の精度を向上させる役割を果たす可能性が示唆される。したがって、今後は 6 月や 7 月の 4 偏波 PALSAR-2 データをアーカイブ以外の方法で確実に取得して分類を行い、精度の向上がみられるか検討することや、これらのデータを RADARSAT-2 データと併用した分類を行うことで、精度の向上を検討することを今後の課題とする。

## 6 Cバンド SAR および光学センサを使用した水稻穂含水率の推定

### 6.1 背景と目的

農業分野において、リモートセンシングに対する情報ニーズとして、農地利用状況の把握、作物生産性の定量化、生育・診断や精密農業管理、災害実体の把握の4点が挙げられる(井上, 2016)。このうち作物生産性の定量化は水稻作でも重要であり、圃場ごとの生育状況は、農作業の適切な管理、収穫適期の予測、品質や収量の予測に重要な指標となる。とくに収穫適期の予測に関しては、これまで出穂日以降の積算気温による方法で判断されてきたが、地域全体ですべての圃場の出穂日を調査することは、時間的、労働力的に非常に困難である(境谷・井上, 2013)。一方、出穂日がわからなくても、収穫適期までの日数がある時点での穂含水率から予測できるため(佐々木・伊藤, 2000)、穂含水率を推定することで圃場間での水稻の収穫適期の早晚を把握することができる。この情報を用いることで、大規模経営や大型機械の共同利用が行われている集団営農等では、各圃場の穂含水率は適切な順番で収穫を実施するための定量的な指標となり、収穫適期の予測日を把握するために有益であると考えられる。しかしながら、穂含水率のような生育指標についても、これを圃場ごとに把握するためには、人の手による現地調査には限界があり、継続的かつ広域的な把握はきわめて困難である。したがって、これまでの現地調査手法の代替としての新たな情報取得の手法を確立することが必要であり、農業地域における課題の一つとなっている。

人工衛星を使用した衛星リモートセンシング技術は、作物生産性の定量化への適用に期待され、特に多くの研究例が報告されている。既往の研究には、光学センサを利用して水稻の生育状況を推定するものが多いといえる。例えば、Landsat データから算出された NDVI などの変数を用いることにより、水稻の草高や被覆率が推定されている(Oguro *et al.*, 2003; Yaghouti *et al.*, 2019)。とくに赤色光や近赤外線波長の波長から算出される NDVI は、計算が簡単なだけでなく、植物の活性に応じて値が変化するため、生育状況の推定に用いられている。しかし、光学センサによる観測では、大気による減衰が大きく、特に煙霧や雲などの天候の影響を受けやすい。そのため、推定精度が大幅に低下するおそれがあり、使用できる衛星データ数が限られてしまう可能性がある(Blaes *et al.*, 2005)。

光学センサのほか、近年では SAR によるマイクロ波を使用した能動型センサも、生育状況の推定に利用されている (Sonobe *et al.*, 2014c)。SAR は地表面の特性に対する強い感度を有しているため、SAR を用いた手法は生育状況の推定に対し信頼できる値を定常的に提供できると考えられる。また、SAR は自身で照射したマイクロ波を観測するため、昼夜を問わず安定した観測が可能であり、大気や天候の影響が少ないデータを取得できる (Bindlish and Barros, 2001; Mascolo *et al.*, 2019)。したがって、毎年安定して確実に観測結果を得られる手法であることから、適切な時期の観測データを使用した生育状況の推定に効果的であると考えられる。SAR を利用した生育状況の推定にも研究例は多く、X バンド SAR を使用した水稻の播種日の推定 (Mascolo *et al.*, 2019) や、水稻のフェノロジーの推定 (Kucuk *et al.*, 2016) などが挙げられる。C バンド SAR では、草高やバイオマス量を推定する研究が多く報告されている (例えば Sharifi and Hosseingholizadeh, 2018; Ndikumana *et al.*, 2018)。また、光学センサと C バンド SAR を併用することにより、水稻の生育状態をより高精度で把握できることも示されている (Torbick *et al.*, 2011; Ramadhani *et al.*, 2020)。このように、水稻の生育状態を C バンド SAR から推定することや、光学センサを用いて精度の向上を図れることの有効性が示されている。

一方、穂含水率を推定する研究として、X バンド SAR を用いた小麦の穂含水率の推定に関する研究例が報告されている (Sonobe *et al.*, 2014c)。しかし、SAR を用いた水稻の穂含水率推定の研究や、C バンド SAR と光学センサを併用した穂含水率を推定する研究、水稻における穂含水率に注目した研究の例はみられない。先述のとおり、穂含水率を推定することで収穫適期の予測ができると考えられ、穂含水率の推定は極めて有用な情報となる。また、SAR データは様々な入射角で観測したものを使用することができるが、水稻の生育状態を推定するにあたり、この入射角の影響を比較検討した研究は、C バンド SAR の入射角ごとのバイオマス量の推定精度を比較した報告 (Valcarce *et al.*, 2019) などわずかな例しかなく、特に穂含水率を推定する研究における検討例はみられない。また、既往の研究の多くが海外での研究事例であるが、現在の北海道でみられるような乾田直播栽培と移植栽培の 2 種類の栽培形態に着目し、水稻の生育状態を推定する研究例は、海外を含めみられない。

本章では、水稻の生育状況を高い精度で推定していくことを目指して、SAR および光学センサを併用して水稻の穂含水率を推定するための検討を行った。SAR には波長帯の異なるいくつかの種類があるが、先述のとおり、水稻の生育状態を C バンド SAR から推定する

ことの有効性や、光学センサを用いることによる精度向上が示されている。作物を対象とした生育状況の推定にはCバンドが適していると考えられるため、本章でもCバンドSAR衛星を使用することとした。また、3章で示したSARを利用した作付作物の分類に関する研究では、CバンドSARデータから算出することができる9種類の変数を用いることにより、高い精度での作付作物分類が可能となった。そこで本章でもこの9変数を算出し、生育状況の推定に適用した。また光学センサは、高解像度衛星の画像を使用し、NDVIの算出を行い、これらを推定に使用した。

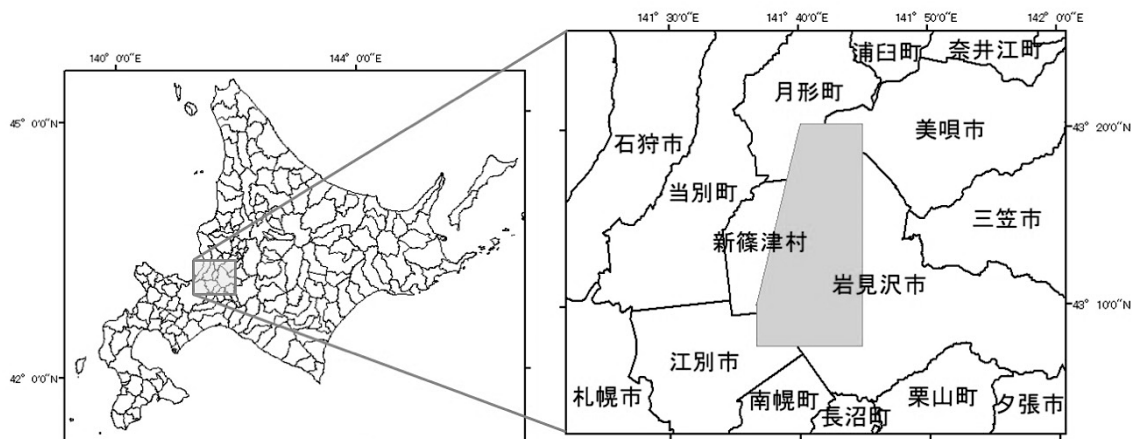
まず、SARデータのみを使用したモデルの有効性を評価するため、水稻における穂含水率の推定モデルを複数作成し、それぞれの精度を比較した。全てのデータを対象としたモデルに加え、衛星進行方向や栽培形態の条件によってデータを分割して作成したモデルを計4種類作成し、それぞれの精度評価からデータ分割法や推定に効果的な変数について検討した。次に、光学センサとSARの双方のデータを併用することの有効性について検討した。光学センサデータのみを使用したモデルと、SARデータおよび光学センサデータを併用したモデルを作成し、精度を比較した。最後に、SARで作成したモデルが翌年以降も使用できるかを判断した。SARデータから作成した5種類のモデルについて、翌年のSARデータを使用して検証を行った。

## 6.2 研究方法

### 6.2.1 対象地域および対象作物

研究対象地域は、北海道空知管内の南部に位置する岩見沢市の北村地域を中心とした稲作地帯（東経  $141^{\circ} 30' 40'' \sim 141^{\circ} 40' 30''$ 、北緯  $43^{\circ} 00' 40'' \sim 43^{\circ} 20' 00''$ ）とした（図 6-1）。石狩川の左岸に沿った石狩平野の下流に位置しているため、対象地域の標高は  $4\text{m} \sim 20\text{m}$  と平坦な地形となっている。

この地域の年平均気温は  $7.6^{\circ}\text{C}$  で、夏季は高温になることもあるが、おおむね冷涼であり、冬季の気温は氷点下となる。年平均降水量は  $1163\text{mm}$  であり、一年を通じて降水がある。夏季から冬季にかけて特に降水量が多く、冬季には積雪があり根雪となる。岩見沢市における



「国土数値情報 (行政区画データ)」 (国土交通省) <<https://nlftp.mlit.go.jp/ksj/index.html>> を加工して作成

図 6-1 対象地域

夏季の日照率を図 6-2 に示す。岩見沢市は夏季の日照率が低く、特に 6 月から 8 月までは 40%を下回っている。したがって、光学センサのみを使用して必要な時期の水稲の穂含水率の情報を得るにはリスクが伴い、SAR を用いて水稲の穂含水率を推定する意義があるといえる。

本地域では主に水稲作が行われているが、水田では転作作物として大豆、小麦、玉葱なども栽培されるなど、空知型輪作体系と呼ばれる田畑輪換による輪作が行われている。なお、

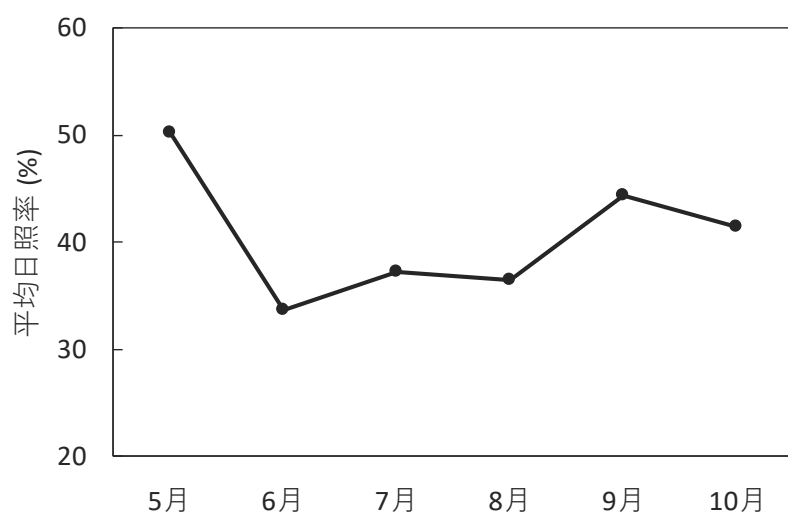


図 6-2 岩見沢市における夏季の平均日照率 (平年値)

本地域の水稲の栽培形態には移植栽培と乾田直播栽培があり，本章では両方の圃場を対象とした。研究は2016年から2019年までの4年間実施しており，各年の対象圃場数を表6-1に示す。各圃場で作付されている品種は，移植栽培がゆめぴりか，ななつぼし，おぼろづき，きらら397の4品種，乾田直播栽培が大地の星，ほしまる，えみまる，そらゆたかの4品種である。

### 6.2.2 使用データおよび処理方法

使用した衛星データは，Cバンドのセンサを搭載したRADARSAT-2の4偏波SARデータおよび高分解能光学センサのGeoeye-1の光学センサデータである。

SARデータは，Fine Quadモードにより観測されたものを使用した。ピクセルスペーシングは8.0m，処理レベルはSingle Look Complex (SLC) である。表6-2にそれぞれの衛星デー

表 6-1 対象圃場数

	移植栽培	乾田直播栽培	合計
2016年	9	4	13
2017年	6	6	12
2018年	6	5	11
2019年	6	5	11

表 6-2 使用した衛星データおよび現地調査日程

年	RADARSAT-2			Geoeye-1	現地調査日
	日付	入射角	衛星進行方向	日付	
2016年	09月03日	FQ11 (30.2~32.0°)	ディセンディング		09月04日
2017年	08月26日	FQ9 (28.0~29.8°)	アセンディング		08月26日
	08月29日	FQ11 (30.2~32.0°)	ディセンディング	09月01日	08月29日
2018年	08月21日	FQ9 (28.0~29.8°)	アセンディング		08月24日
	08月24日	FQ11 (30.2~32.0°)	ディセンディング		
2019年	08月20日	FQ11 (30.2~32.0°)	ディセンディング		08月20日
	08月26日	FQ19 (38.3~39.8°)	アセンディング		08月27日

データの観測時期，入射角，衛星進行方向を示す。RADARSAT-2 では観測時のビーム照射角度に付された番号が入射角に対応し，FQ とそれに続く数字で表される。本章では，入射角が約 30°の FQ9，FQ11 と，入射角が約 40°の FQ19 を使用した。FQ9 と FQ19 は上昇軌道（アセンディング）で観測されており，FQ11 は下降軌道（ディセンディング）で観測されている。なお，それぞれの衛星データは出穂期から成熟期の間で観測されたものを使用したため，各年とも 8 月下旬から 9 月上旬までのものとなっている。

Geoeye-1 の光学センサデータにはパナクロマチックデータとマルチスペクトルデータがあるが，本章ではマルチスペクトルデータを使用した。ピクセルスペーシングは 1.64m である。光学センサデータは SAR データの観測に近い時期のデータをすべて使用する予定であったが，気象条件が悪く，観測できたのは 2017 年 9 月 1 日の 1 回のみであった。

推定モデルを作成するにあたり，RADARSAT-2 データの前処理には，前章までと同様にノイズ除去，マルチルック処理，オルソ補正を行った。光学センサデータの前処理には，FLAASH 法による大気補正を行って反射率に変換し，SAR データと同様にオルソ補正を行った。これらの処理には，Harris Geospatial Solutions 社製ソフトウェア ENVI 5.5+IDL 8.7 を使用した。

以上のように調整したデータを用いて，モデルに使用する変数を算出した。SAR データからは，4 章と同様に HH，HV，VV 偏波の $\sigma^0$ による後方散乱係数と，Van Zyl による散乱モデル分解法の散乱成分 (van Zyl *et al.*, 2011) の 2 回散乱 (Pd)，体積散乱 (Pv)，表面散乱 (Ps)，Cloude-Pottier による固有値解析法の散乱成分 (Cloude and Pottier, 1996) のエントロピ (H)，アニソトロピ (A)，アルファ角 ( $\alpha$ ) の計 9 変数を算出した。

光学センサデータからは，赤色光のバンド 3 (655–690 nm) と近赤外光のバンド 4 (780–920 nm) から，(6-1) 式より NDVI を算出し，この 1 変数を解析に使用した。

$$\text{NDVI} = \frac{\text{Band4} - \text{Band3}}{\text{Band4} + \text{Band3}} \quad (6-1)$$

ここで，Band3 はバンド 3，近赤外光のバンド 4 の反射率である。

### 6.2.3 解析方法

作成した推定モデルは，SAR の 9 変数および光学センサの 1 変数を説明変数とし，穂含

水率を目的変数として作成した。穂含水率を算出するため、各圃場で現地調査を行った。現地調査は SAR の観測時期に合わせて行っている (表 6-2)。各圃場につき 3 地点で 0.25m<sup>2</sup> 区画の坪刈りを行い、得られた水稻のうち 300g を無作為に抽出した。この水稻を穂と茎葉に切り分け、穂の新鮮重量を測定した。穂は通風式乾燥機に入れ、60°C で 48 時間乾燥させたのち、乾燥重量を測定した。このようにして測定した乾燥前後の質量をもとに、以下の式 (6-2) で穂含水率の算出を行った。

$$MCE = \frac{FWE - DWE}{FWE} \quad (6-2)$$

ここで、MCE (Moisture Content of Ear) は穂含水率、FWE (Fresh Weight of Ear; kg) は穂の新鮮重量、DWE (Dry Weight of Ear; kg) は穂の乾燥重量である。

目的変数である穂含水率は圃場単位の値としているため、説明変数も圃場単位で値を算出して推定モデルを作成することとした。まず、各年の調査圃場の境界を示す GIS データを作成した。この GIS データを用いて、各衛星データから算出した全ての説明変数について圃場ごとにバッファの除去や平均値の算出を行った。

光学センサデータのみを用いたモデル作成では単回帰分析を行い、それ以外のモデル作成では重回帰分析を行った。SAR のみを用いたモデルでは、重回帰分析にステップワイズ法の一つである変数減少法を用いて、説明変数の選択を行った。変数減少法の評価指標には赤池情報量基準 (AIC) を使用した。次に、選択されたすべての変数の VIF 統計量を算出した。VIF 統計量は、重回帰分析において各変数に多重共線性がないかを調べるための指標である。算出した VIF 統計量が 5 を超えている変数は除き、残りの変数を用いてモデルを作成した。SAR と光学センサを併用したモデルでは、データ数が少なくステップワイズ法を用いることができなかつたため、変数の選択には総当たり法を使用した。AIC が最も小さく、かつ全ての変数で VIF 統計量が 5 を超えていない条件で変数を選択し、重回帰分析によるモデルを作成した。

作成したモデルの精度評価の指標として、単回帰分析では決定係数 (R<sup>2</sup>)、重回帰分析では自由度調整済み決定係数 (R<sup>2</sup>) および二乗平均平方根誤差 (RMSE) も併せて使用した。ステップワイズ法によって選択された変数については、標準偏回帰係数を算出してモデルへの寄与度を評価した。本章では作成したモデルの精度検証も行ったが、その精度評価には

RMSE を使用した。なお、これらのモデル作成および検証には、フリー統計ソフトである R (R Core Team, 2016) を使用した。

#### 6.2.4 解析の流れ

本章では、以下の 1) ~ 3) で示す 3 処理で解析を進めた。

##### 1) SAR データを使用した水稻穂含水率の推定モデル作成 (6.3.2 参照)

まず、2016 年から 2018 年までの計 5 シーンの SAR データを使用して、9 変数からすべての圃場を対象とした水稻の穂含水率の推定モデルを作成した。作成されたモデルについて決定係数と RMSE を用いた精度の評価を行い、外れ値となった圃場について考察した。

次に、2016 年から 2018 年までの SAR データのうち、アセンディング軌道で観測された FQ9 データ (計 2 シーン) またはディセンディング軌道で観測された FQ11 データ (計 3 シーン) のいずれかのみを使用し、衛星進行方向ごとに推定モデルを作成した。これらのモデルについても、SAR の 9 変数を用い、全ての圃場を対象とした。作成された双方のモデルについて精度の評価を行い、全てのデータを使用したモデルとの間で精度の比較を行った。

最後に、2016 年から 2018 年までの計 5 シーンの SAR データを使用して、9 変数から栽培形態ごとに水稻の穂含水率の推定モデルを作成した。直播栽培用のモデルと移植栽培用のモデルについて両者の精度の評価を行い、全てのデータを使用したモデルとの間で精度の比較を行った。

##### 2) SAR データおよび光学センサデータを使用した水稻穂含水率の推定モデル作成 (6.3.3 参照)

まず、2017 年の 1 シーンの光学センサデータを使用して、NDVI からすべての圃場を対象とした水稻の穂含水率の推定モデルを作成し、精度の評価を行った。

次に、2017 年の 1 シーンの SAR データに光学センサデータを併用して、計 10 変数からすべての圃場を対象とした水稻の穂含水率の推定モデルを作成した。作成されたモデルについて精度の評価を行い、SAR データと光学センサデータを併用することによる精度の変化を検討した。

##### 3) SAR データを使用した水稻穂含水率の推定モデル検証 (6.3.4 参照)

1) で作成した 5 つのモデル (全データモデル、アセンディング軌道によるモデル、ディセンディング軌道によるモデル、直播栽培モデル、移植栽培モデル) について、2019 年の 2

シーンのデータを用いて検証を行った。全データモデルについては、2019年8月20日と8月26日のデータを適用し、それぞれのRMSEを算出し、作成したモデルの汎用性について検討した。同様にアセンディング軌道によるモデルは8月26日、ディセンディング軌道によるモデルは同年8月20日のデータを適用して精度を検証した。直播栽培モデル、移植栽培モデルに関しても、同様に同年8月20日と8月26日のデータのうち直播栽培のみ、移植栽培のみをそれぞれ適用して精度を検証した。

### 6.3 結果と考察

#### 6.3.1 現地調査結果

各年各圃場の穂含水率の実測結果を図6-3～図6-6に示す。

2016年の現地調査は9月4日の1回のみである(図6-3)。調査時期が遅かったため、全体的に30%～45%と他の年よりも含水率が低い値となった。移植栽培Dは44.4%であったが、それ以外の移植栽培は平均35.6%、直播栽培は平均37.1%で、移植栽培の方が低い含水率となっている傾向がみられた。

2017年の現地調査は8月26日と8月29日の2回行った(図6-4)。2回とも移植栽培の含水率は30%～45%、直播栽培の含水率は45%～60%と、栽培形態ごとに含水率の違いがみられた。また、移植栽培の1圃場を除き、2回目の調査で含水率の低下がみられた。多くの圃場で、8月26日から8月29日の3日間で3ポイント程度低下していた。

2018年の現地調査は8月24日の1回のみである(図6-5)。現地調査を行った日が8月24日と過去2年よりやや早く、穂含水率は移植栽培で40%～60%、直播栽培では55%～65%と高い値を示した。

2019年の現地調査は2回行った(図6-6)。8月20日については、直播栽培と移植栽培で1圃場ずつ降雨のために調査を中止したため、欠測となっている。2019年は使用した衛星データの片方が2017年や2018年と異なるFQ19であるために、8月20日、8月27日と2回の調査日の間隔が7日間離れている。2019年は各調査日で直播栽培、移植栽培で明確な差がなかった。とくに直播栽培において過去3年より低い値となっていた圃場が散見された。すべての圃場において、2回目の調査で穂含水率は低下していた。また、7日間で概ね5～

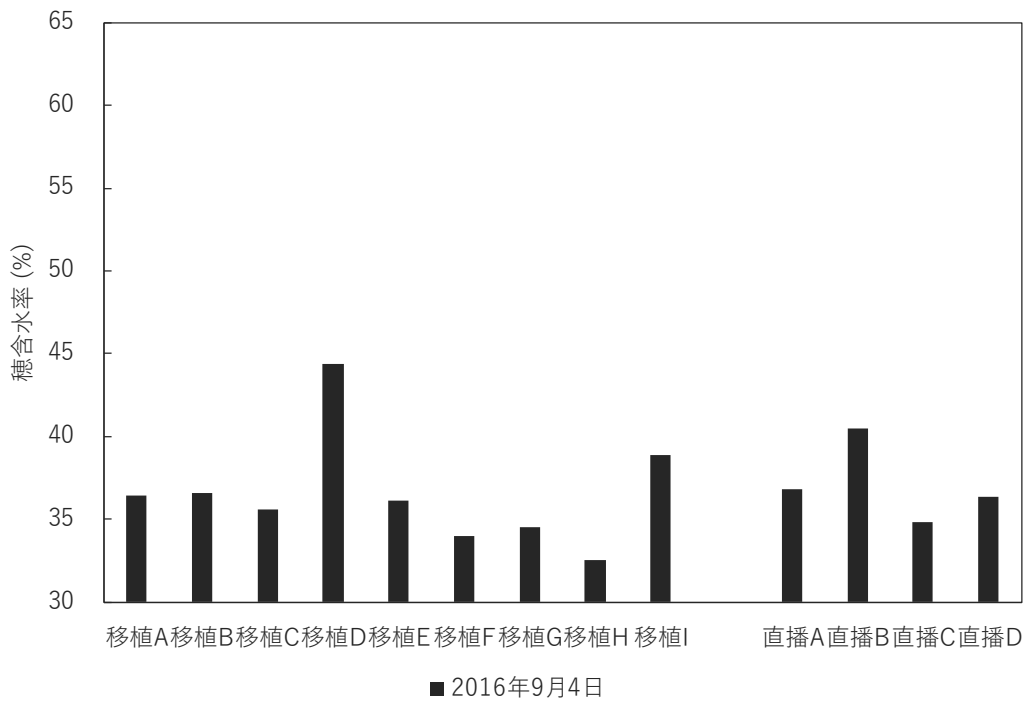


図 6-3 2016 年の各圃場における穂含水率

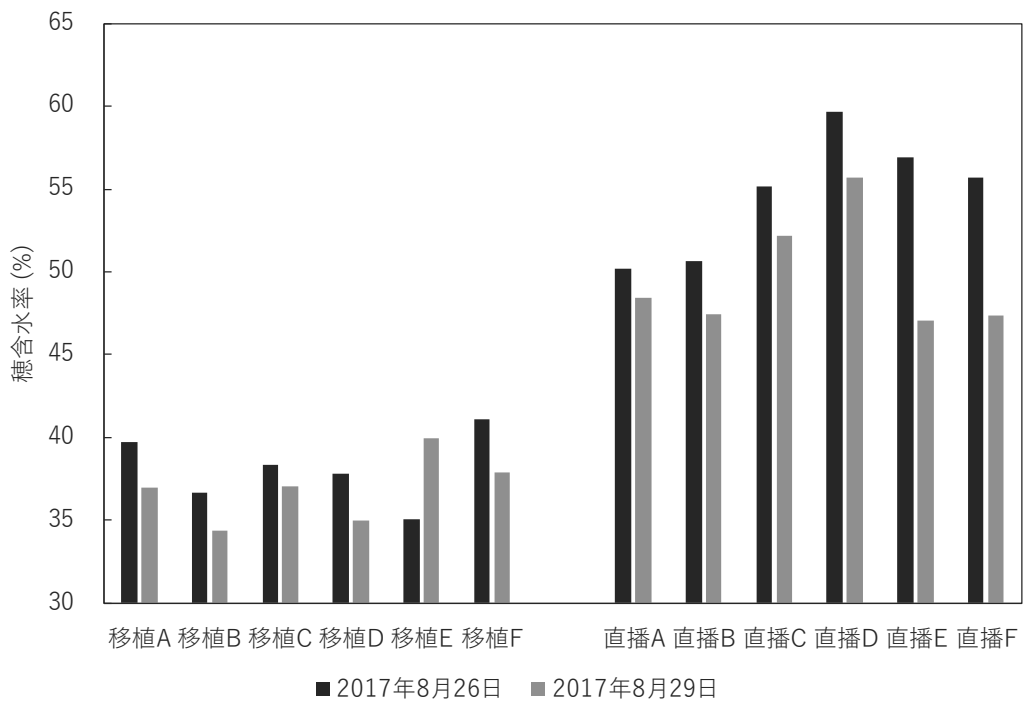


図 6-4 2017 年の各圃場における穂含水率

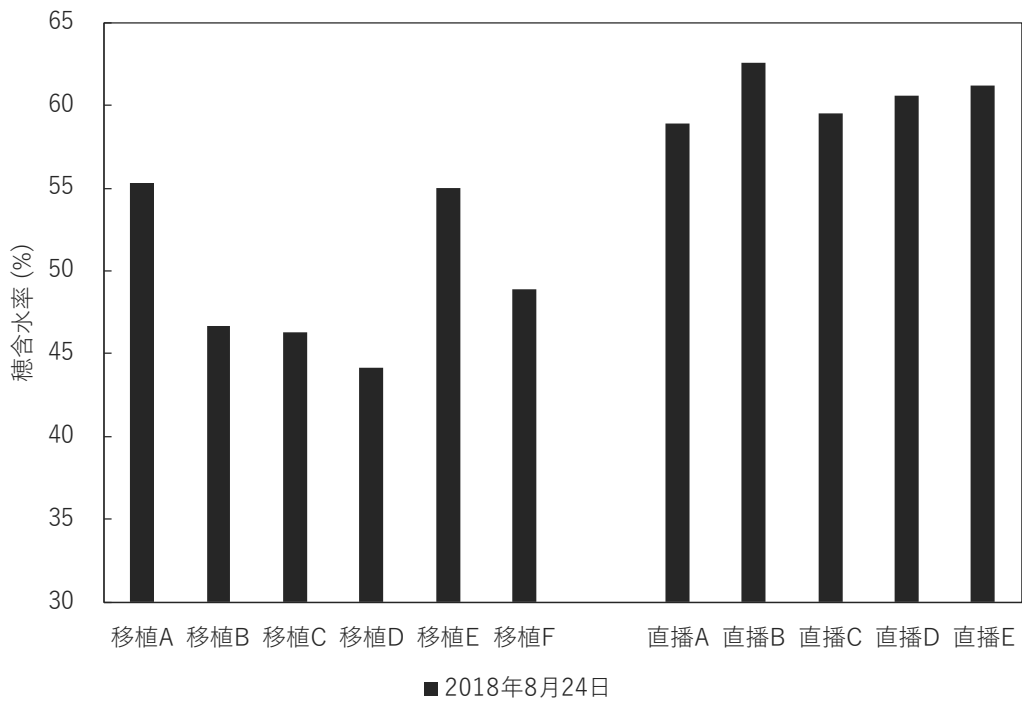


図 6-5 2018 年の各圃場における穂含水率

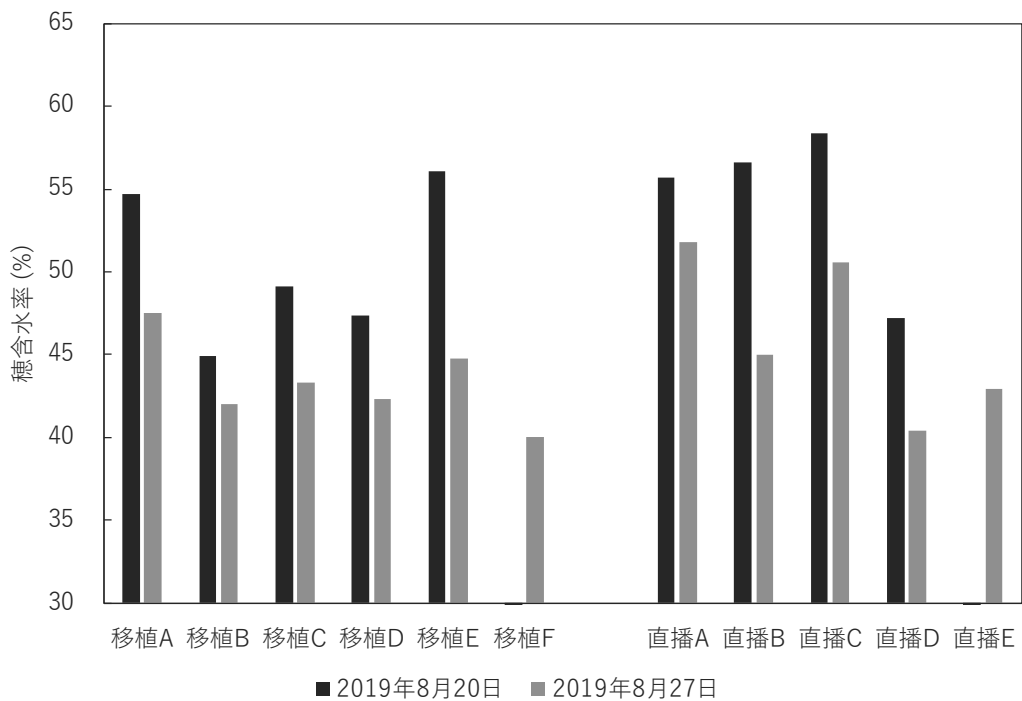


図 6-6 2019 年の各圃場における穂含水率

10 ポイント程度の低下がみられた。

### 6.3.2 SAR データを使用した水稻穂含水率の推定モデル作成

#### 1) 全データによるモデルの作成

2016 年から 2018 年までの 3 年間に観測されたすべての SAR データを用いたモデルを作成し、実測値と推定値を比較した結果を図 6-7 に示す。決定係数は 0.152 となっており、全てのデータを用いたモデルでは穂含水率を十分には説明することはできなかった。実測値が高い圃場では推定値が過小評価され、逆に実測値が低い圃場では過大評価される傾向がみられた。

作成されたモデルを式 (6-3) に示す。

$$\text{MCE} = 2.340 \times \text{VV} - 43.659 \times \text{H} + 107.058 \quad (6-3)$$

モデルには VV 偏波後方散乱係数、エントロピの 2 変数が選択された。これらの変数について標準偏回帰係数を示したのが表 6-3 である。VV 偏波後方散乱係数は標準偏回帰係数の絶対値が大きく、穂含水率の推定への影響が大きいといえる。先行研究でも、VV 偏波は HH

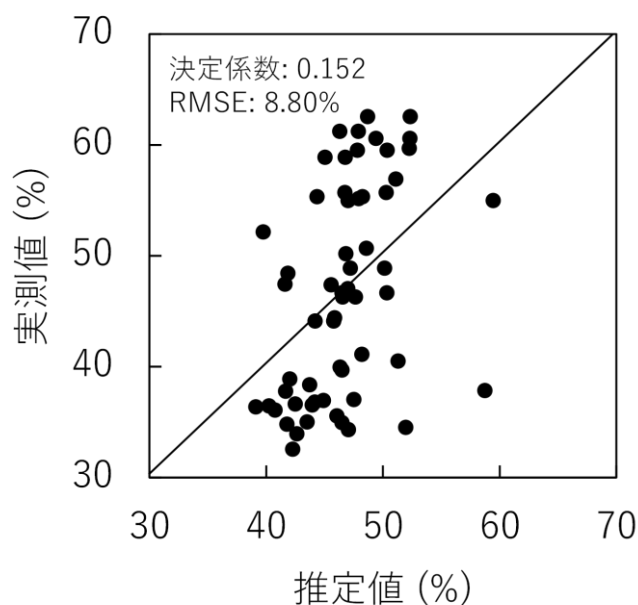


図 6-7 全データを使用したモデルによる穂含水率の推定値と実測値の関係

表 6-3 全データを使用したモデルに選択された各変数の標準偏回帰係数

変数	標準偏回帰係数
VV	0.323
H	-0.277

偏波より作物体に関する情報の収集に適しているとされており (Englhart *et al.*, 2011), 本研究も同様の結果となったと考えられる。エントロピは散乱のランダム性を表す変数であるため、穂含水率の多寡によって変化したと考えられる。

## 2) 衛星進行方向別でのモデルの作成

2016年から2018年までの3年間に観測された SAR データのうち、ディセンディング軌道で観測された FQ11 データ (計3シーン) のみを使用して作成したモデルと、アセンディング軌道で観測された FQ9 データ (計2シーン) のみを使用して作成したモデルを作成し、実測値と推定値を比較した結果を図 6-8 に示す。ディセンディング軌道とアセンディング軌道によるモデルの決定係数はそれぞれ 0.263, 0.661 となっており、衛星進行方向でモデルを分けることにより全データを使用したモデルと比較して高い精度を得ることができた。以上の結果から、衛星進行方向が穂含水率を推定する精度に影響を及ぼす可能性が示唆され

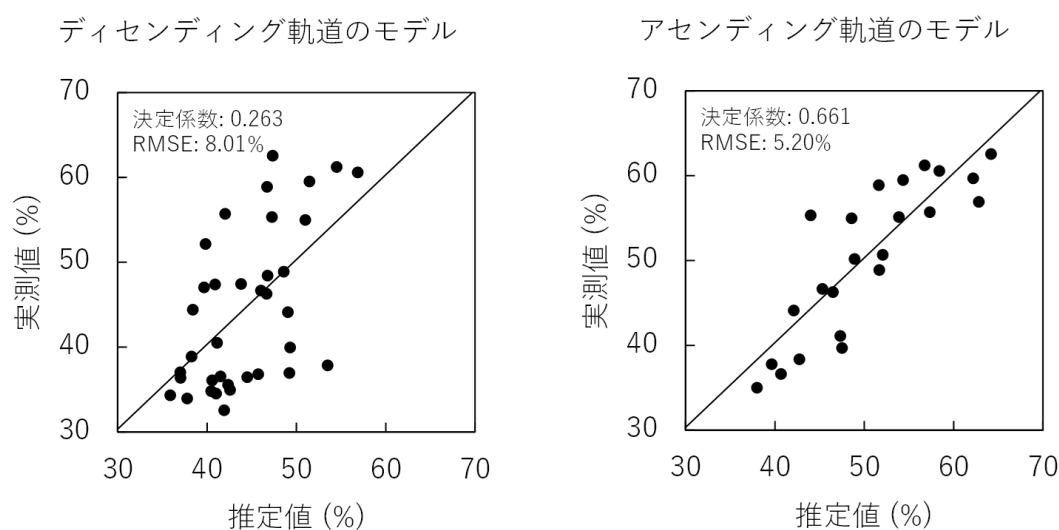


図 6-8 衛星進行方向ごとに作成したモデルによる穂含水率の推定値と実測値の関係

た。とくにアセンディング軌道によるモデルは決定係数が高く、本地域においてはアセンディング軌道による SAR データを使用することで特に水稻の穂含水率を推定しやすいと考えられる。この理由として、ディセンディング軌道とアセンディング軌道でデータ観測時間が異なっていることが考えられる。本地域では、ディセンディング軌道の観測は早朝、アセンディング軌道の観測は夕刻に行われているため、ディセンディング軌道の観測時には朝露が発生している可能性が考えられる。ディセンディング軌道によるモデルは、全データを使用したモデルと同様に実測値が高い圃場では推定値が低く、過小評価されている傾向がみられ、実測値が低い圃場では推定値が高く、過大評価されている傾向がみられた。したがって、ディセンディング軌道による観測時間に発生した早朝の朝露が、レーダの散乱に影響を与え、精度の低下を及ぼしていることが考えられる。

ディセンディング軌道およびアセンディング軌道によるモデルを、それぞれ式 (6-4)、式 (6-5) に示す。

$$\text{MCE} = 1.155 \times \text{Ps} + 79.037 \times \text{H} - 1.060 \times \text{Alp} - 40.871 \quad (6-4)$$

$$\begin{aligned} \text{MCE} = 0.723 \times \text{Pd} + 3.641 \times \text{Pv} - 97.156 \times \text{H} + 82.319 \times \text{A} \\ - 202.006 \end{aligned} \quad (6-5)$$

ディセンディング軌道のモデルには表面散乱、エントロピ、アルファ角の 3 変数が選択され、アセンディング軌道のモデルには 2 回散乱、体積散乱、エントロピ、アニソトロピの 4 変数が選択された。これらの変数について、目的変数の推定に寄与している変数を標準偏回帰係数で比較した (表 6-4, 6-5)。ディセンディング軌道については、エントロピやアルファ

表 6-4 ディセンディング軌道のデータのみを使用したモデルに選択された  
各変数の標準偏回帰係数

変数	標準偏回帰係数
Ps	0.371
H	0.500
Alp	-0.459

表 6-5 アセンディング軌道のデータのみを使用したモデルに選択された

各変数の標準偏回帰係数	
変数	標準偏回帰係数
Pd	0.161
Pv	0.389
H	-0.697
A	0.409

角の標準偏回帰係数の絶対値が大きく、穂含水率の推定に影響していた。アルファ角は反射角の平均を表す数値であり、穂の水分の差異によって反射の動向が異なると考えられ、それによって穂含水率の推定に有効な変数であったと考えられる。アセンディング軌道については、エントロピで特に標準偏回帰係数の絶対値が大きく、-0.6 を下回っていた。その他、アニソトロピ、体積散乱が標準偏回帰係数の絶対値が大きい値として続いた。この2変数は作物における被植の繁茂の程度に応じて値が変化しやすいと考えられる。体積散乱は作物体が大きくなることで起こりやすくなる散乱であるため、推定モデルに効果的な変数であるといえる。全データのモデルや衛星進行方向ごとのモデルでは、エントロピがすべてに共通して選択されていたため、特に重要であると考えられる。

### 3) 栽培形態別でのモデルの作成

2016年から2018年までの3年間に観測された全5シーンのSARデータを使用し、移植栽培、直播栽培のみをそれぞれ対象として水稻の穂含水率の推定モデルを作成した。これらのモデルについて、実測値と推定値を比較した結果を図6-9に示す。移植栽培のモデルと直播栽培のモデルの決定係数はそれぞれ0.528、0.308となっており、栽培形態でモデルを分けることにより全データを使用したモデルと比較して高い精度を得ることができた。以上の結果から、栽培形態が穂含水率を推定する精度に影響を及ぼす可能性が示唆された。この理由として、直播栽培と移植栽培では登熟の時期や収穫の時期に1週間程度の差がある点や、条幅が異なる点などが考えられる。また、直播栽培は他年より衛星観測時期の遅い2016年のデータで外れ値が多くみられたため、登熟が進むことで精度の低下がみられることも示唆される。2016年以外のデータについては、栽培形態によってモデルを分けることにより、これまでのモデルでみられていた高い含水率の圃場の過小評価や低い含水率の圃場の過大

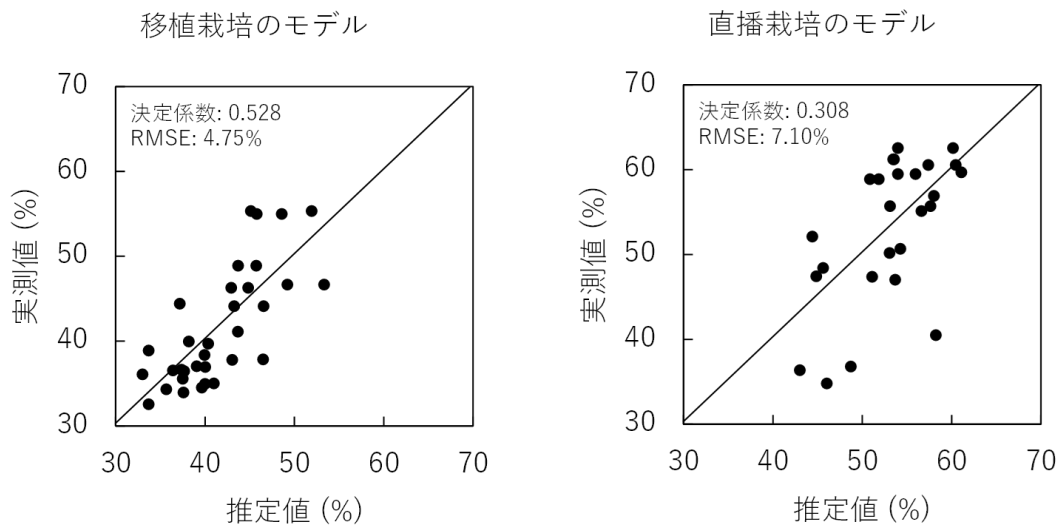


図 6-9 栽培形態ごとに作成したモデルによる穂含水率の推定値と実測値の関係

評価がみられにくくなった。すなわち、移植栽培と直播栽培でモデルを分けることにより、安定して高い精度でモデルの作成がなされる可能性が示唆された。

移植栽培のモデルおよび直播栽培のモデルを、それぞれ式 (6-6)、式 (6-7) に示す。

$$\text{MCE} = 3.515 \times \text{HV} + 0.511 \times \text{Ps} + 69.922 \quad (6-6)$$

$$\text{MCE} = 2.685 \times \text{VV} - 72.460 \times \text{H} + 137.940 \quad (6-7)$$

移植栽培のモデルには HV 偏波後方散乱係数, 表面散乱の 2 変数が選択され, 直播栽培のモデルには VV 偏波後方散乱係数, エントロピの 2 変数が選択された。これらの変数について, 目的変数の推定に寄与している変数を標準偏回帰係数で比較した (表 6-6, 6-7)。移植栽培については, HV 偏波後方散乱係数で標準偏回帰係数が 0.6 を上回っており, 穂含水率の

表 6-6 移植栽培のみを使用したモデルに選択された各変数の標準偏回帰係数

変数	標準偏回帰係数
HV	0.640
Ps	0.206

表 6-7 直播栽培のみを使用したモデルに選択された各変数の標準偏回帰係数

変数	標準偏回帰係数
VV	0.414
H	-0.426

推定に影響していた。HV 偏波は作物体の大きさによって値が大きくなるため、体積散乱と同様に作物の大きさを捉える役割を果たしたと考えられる。直播栽培については、VV 偏波後方散乱係数とエントロピがそれぞれ 0.4, -0.4 程度と同等の標準偏回帰係数を示していた。本結果においても、これまでの結果と同様に、体積散乱や HV 偏波後方散乱係数、VV 偏波後方散乱係数など、作物の大小や形状に影響する変数や、エントロピが選択され、かつ高い標準偏回帰係数を示してており、これらの変数の有用性が示される結果となった。

### 6.3.3 SAR データおよび光学センサデータを併用した水稻穂含水率の推定モデル作成

#### 1) 光学センサデータによるモデルの作成

2017 年の 1 シーンの光学センサデータから算出された NDVI を使用して穂含水率の推定モデルを作成し、実測値と推定値を比較した結果を図 6-10 に示す。決定係数は 0.413 となっており、良好な結果を得ることができなかった。この理由として、1 点の外れ値がみられたことが原因として考えられる。この 1 圃場については、8 月 29 日にもかかわらず唯一湛水がなされており、湛水の影響によって NDVI が低く算定され、過小評価されたことが原因として考えられる。そこで、この 1 点を除き、再度推定モデルを作成した。実測値と推定値を比較した結果を図 6-11 に示す。決定係数は 0.923 となり、大幅な精度の改善がみられた。以上の結果から、高解像度な光学センサデータの NDVI を用いることにより、高精度な水分推定モデルの作成が可能であることが示された。一方、湛水された圃場があればその影響を受けるために、過小評価されやすくなる可能性があることから、すべての条件を 1 つの式にまとめた NDVI による穂含水率推定は困難である可能性も示された。1 点の外れ値を除いた光学センサデータによるモデルが式 (6-8) である。

$$MCE = 214.58 \times NDVI - 123.84 \quad (6-8)$$

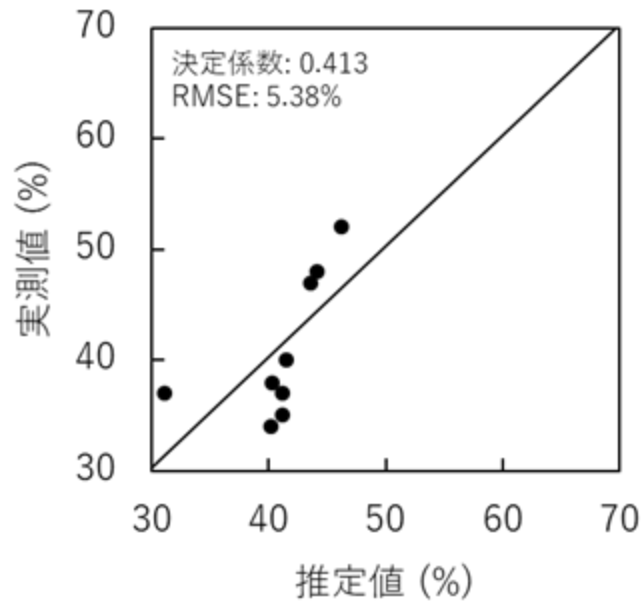


図 6-10 全ての光学センサデータを使用したモデルによる穂含水率の推定値と実測値の関係

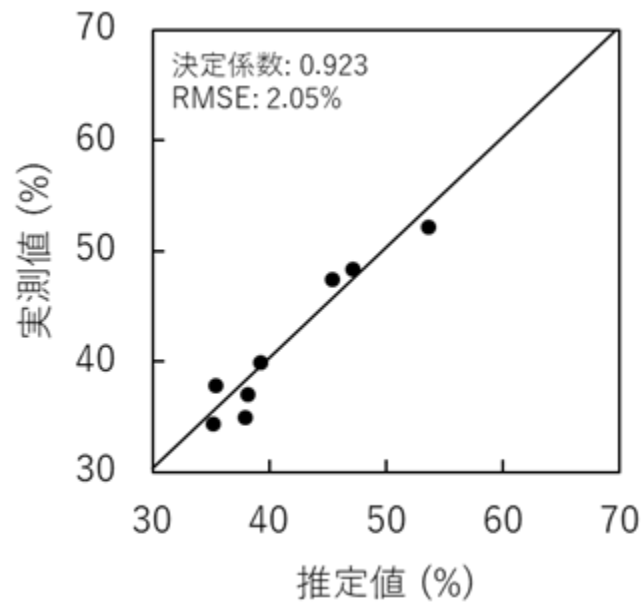


図 6-11 湛水圃場のデータを除いた光学センサデータを使用したモデルによる穂含水率の推定値と実測値の関係

## 2) SAR データと光学センサデータを併用したモデルの作成

2017年の1シーンのSARデータから算出された9変数と、光学センサデータより算出されたNDVIの計10変数を使用し、穂含水率の推定モデルを作成した。先述のとおり、この解析のみデータ数が少なくステップワイズ法を用いることができなかつたため、重回帰分析には総当たり法を使用した。実測値と推定値を比較した結果を図6-12に示す。決定係数は0.943となり、本章で最も良好な結果を得られた。とくに1)において湛水のあった1圃場で外れ値がみられる結果となっていたが、ここではこの1圃場も同じ式で推定することが可能であった。以上のことから、SARの変数は光学センサの変数との併用において、土壌や水面などの情報を補正するため、より高精度なモデルの作成に寄与することが示された。

RADARSAT-2データおよびGeoeye-1データを併用したモデルを、それぞれ式(6-9)に示す。

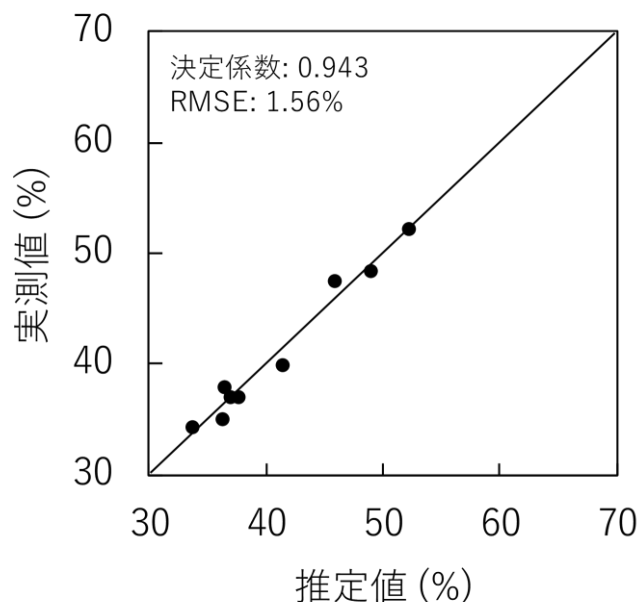


図6-12 RADARSAT-2データおよびGeoeye-1データを併用したモデルによる穂含水率の推定値と実測値の関係

$$\begin{aligned} \text{MCE} = & -3.555 \times \text{Pd} - 1.375 \times \text{Ps} + 59.809 \times \text{A} + 52.614 \times \text{NDVI} \\ & + 283.688 \end{aligned} \quad (6-9)$$

モデルには、2回散乱、表面散乱、アニソトロピ、NDVIが選択された。これらの変数について、目的変数の推定に寄与している変数を標準偏回帰係数で比較した(表6-8)。4変数のうち、特に2回散乱は標準偏回帰係数がそれぞれ-0.9を下回ったほか、NDVIは0.4を上回り、表面散乱は-0.4を下回った。NDVIによって穂含水率を直接推定することができていたことから、他のモデルでよく用いられていた体積散乱やエントロピなどは選択されなかった。代わりに、2回散乱や表面散乱、アニソトロピといった変数が選択された。被植の情報をNDVIで、穂が垂れてきた状態や土壌表面の状態など、被植以外の情報をSARの変数でそれぞれ説明し、より精度が向上したと考えられる。

### 6.3.4 SARデータを使用した水稻穂含水率の推定モデルの検証

#### 1) 全データによるモデルの検証

6.3.2の1)で作成した2016年から2018年のSARの全データを使用したモデルについて、2019年の2シーンのSARデータを用いて検証を行った。実測値と推定値を比較した結果を図6-13に示す。8月26日のデータについては、5%程度のRMSEで推定することが可能であった。また、推定値が55%を超えていた1点を除くと、RMSEは3.29%となり、より低いRMSEが得られた。この1点に関しては、特にエントロピの値が他の圃場と比較して低い値を示していた。一方、8月20日のデータについては、RMSEが6.89%であり、8月26日のデータと比較して高いRMSEとなった。以上の結果から、全データモデルについて、作成

表6-8 RADARSAT-2データおよびGeoeye-1データを併用したモデルに

選択された各変数の標準偏回帰係数

変数	標準偏回帰係数
Pd	-0.901
Ps	-0.464
A	0.394
NDVI	0.487

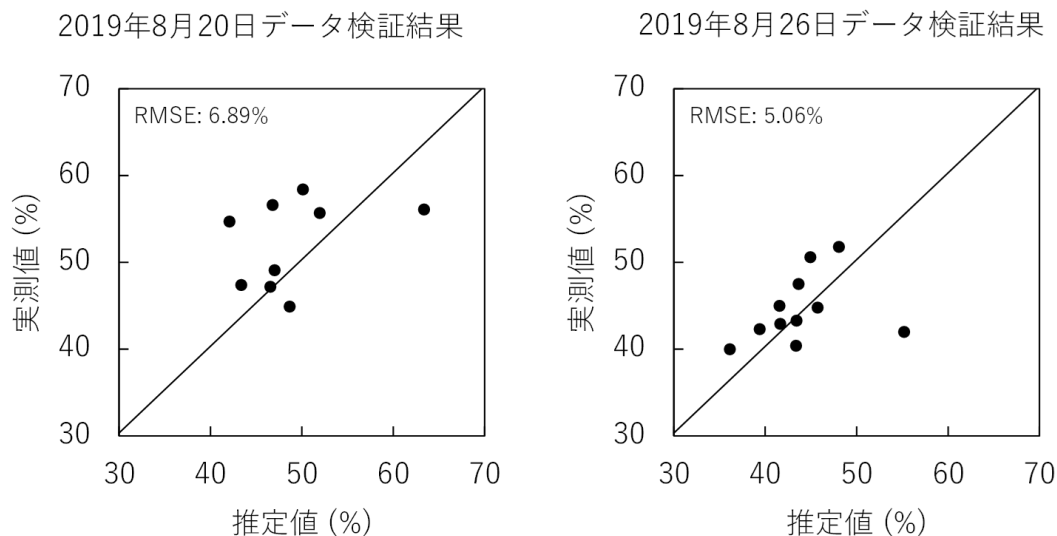


図 6-13 SAR の全データを使用したモデルによる穂含水率の検証結果

した時点での精度に関しては他のモデルよりも低かったものの、推定に関しては 6%程度の RMSE で推定できることから、汎用的なモデルであることが示唆された。

## 2) 衛星進行方向別でのモデルの検証

6.3.2 の 2) で作成した 2016 年から 2018 年までのディセンディング軌道で観測された FQ11 の SAR データを使用したモデルと、アセンディング軌道で観測された FQ9 の SAR データを使用して作成したモデルについて、2019 年の SAR データを用いて検証を行った。ディセンディング軌道によるモデルについては 8 月 20 日 (FQ11)、アセンディング軌道によるモデルについては 8 月 26 日 (FQ19) のデータのみを検証に用いた。実測値と推定値を比較した結果を図 6-14 に示す。ディセンディング軌道によるモデルに関しては、RMSE が 14.21% と全データモデルの 6.89% より高く、多くの圃場で過小評価される傾向がみられた。ディセンディング軌道によるモデルは、3.2.2 で示したとおりモデル自体の精度が低く、検証での精度も低くなったと考えられる。

アセンディング軌道によるモデルに関しては、RMSE が 12.78% と全データモデルの 5.06% より大幅に高く、高い精度での推定が不可能であることが示された。大幅に過小評価されている圃場や、大幅に過大評価されている圃場が多くみられた。8 月 26 日のデータに関しては、アセンディング軌道によるモデルを作成した際に使用した FQ9 のデータより、約 10°入

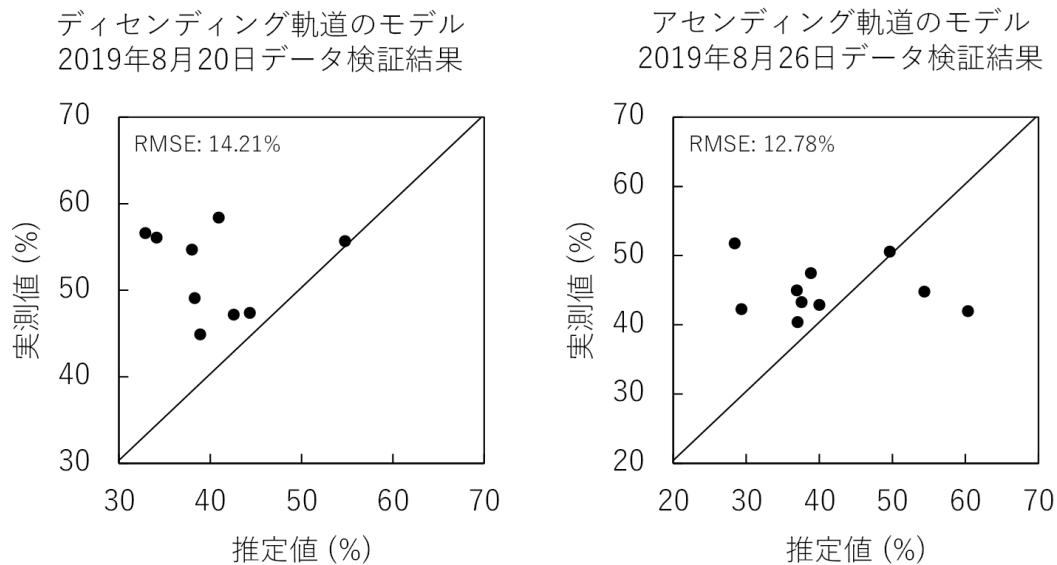


図 6-14 SAR の衛星進行方向ごとに作成したモデルによる穂含水率の検証結果

射角の大きい FQ19 のデータであった。したがって、入射角が異なったことで散乱の形態が異なり、精度の低下に繋がった可能性が考えられる。本章では衛星進行方向についてのみ検討したモデルの作成を行ったが、衛星の入射角についても検討していく必要があると考えられる。

以上の結果から、衛星進行方向ごとに作成した穂含水率の推定モデルに関しては、作成した時点での精度は高かったものの、汎用性の高いモデルではなく、精度の向上を図る必要があると考えられる。とくにアセンディング軌道によるモデルは同じ入射角でのデータで再検証を行うか、FQ19 のデータを含めて FQ19 にも対応したモデルを作成することが必要であると考えられる。

### 3) 栽培形態別でのモデルの検証

6.3.2 の 3) で作成した 2016 年から 2018 年までの 5 シーンの SAR データのうち、移植栽培のみで作成したモデルと、直播栽培のみで作成したモデルについて、2019 年の SAR データを用いて検証を行った。いずれのモデルに関しても、2019 年 8 月 20 日のデータおよび 8 月 26 日のデータを検証に用いた。実測値と推定値を比較した各結果を図 6-15 に示す。直播栽培モデルの RMSE は 3.57% と特に良好な精度が得られた。一方、移植栽培モデルの RMSE

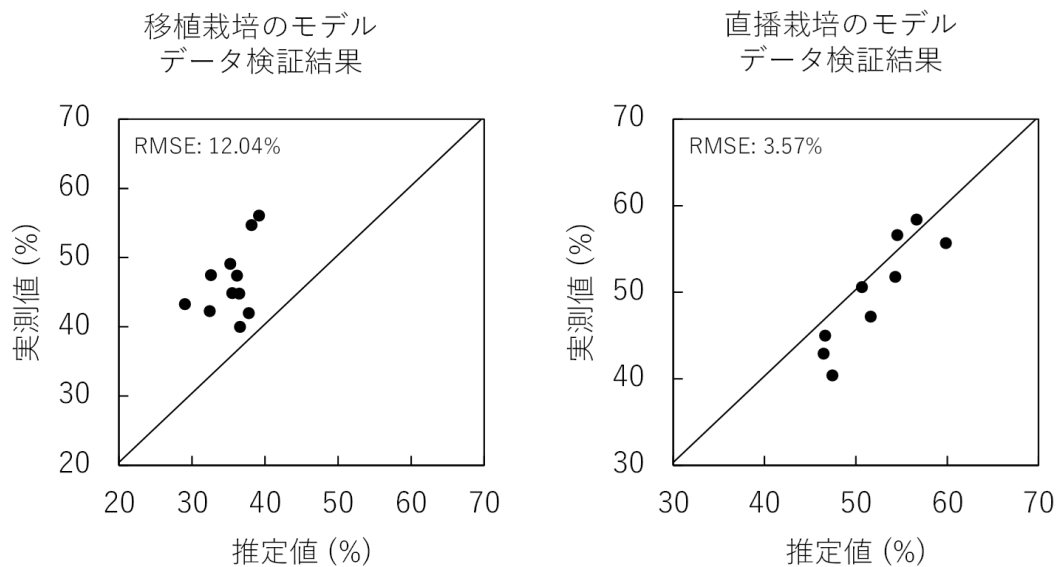


図 6-15 SAR の栽培形態ごとに作成したモデルによる穂含水率の検証結果

は 12%を上回った。この値は全データモデルの 5.06%, 6.89%より高く、良好な精度は得られなかったといえる。この理由として、特に HV 偏波後方散乱係数の値がモデルを作成したデータと比較して低い値を取っているケースがみられた。6.3.4 の 2)で示したように、8月20日のデータはモデル作成の時点で精度の低いディセンディング軌道のデータであり、8月26日のデータは入射角が他のアセンディング軌道のデータより大きいデータであった。そのため、8月20日のデータであれば朝露の影響が、8月26日のデータであれば入射角の違いによる散乱形態の違いの影響が出やすく、精度の低下を招いた可能性が示唆される。また、この影響は直播栽培によるモデルで使用されていない HV 偏波後方散乱係数の値において顕著にみられることが示唆された。以上の結果から、栽培形態別でモデルを作成したところ、直播栽培によるモデルは検証精度が高く汎用性の高いモデルであった。一方で移植栽培によるモデルは作成したモデルの精度は全データによるモデルより高い値が得られたものの、検証の精度は低く、過適合が生じていた可能性がある。移植栽培のモデルは汎用性の高いモデルではなく、更に精度の向上を図る必要があると考えられる。

## 6.4 まとめ

本章では、Cバンド4偏波SARであるRADARSAT-2データと、高解像度光学センサであるGeosyde-1データを併用し、稲作地帯における水稻の穂含水率の推定モデルの作成と検証を行った。

まず、SARデータのみを使用した推定モデルの作成を行った。2016年から2018年までの計5シーンのSARデータを使用して、9変数から推定モデルを作成したところ、高い決定係数を得られなかった。実測値が高い圃場では推定値が低く、過小評価されている傾向がみられ、実測値が低い圃場では推定値が高く、過大評価されている傾向がみられた。

そこで、これら5シーンのSARデータのうち、アセンディング軌道で観測されたFQ9データまたはディセンディング軌道で観測されたFQ11データのいずれかのみを使用し、衛星進行方向ごとに推定モデルを作成した。その結果、それぞれのモデルで全データモデルよりも高い決定係数が得られた。とくにアセンディング軌道によるモデルでは決定係数が0.6を上回り、大幅に精度の向上がみられた。この結果から、モデルの作成には衛星進行方向ごとにデータを分割することで、より精度の向上が期待できることが示された。この原因として、ディセンディング軌道の観測時間が早朝で朝露の影響を受けやすい点などが精度の低下を招く原因として考えられる。したがって、衛星進行方向や観測時間は重要な条件となる可能性が示唆される。

衛星進行方向ごとのモデルと同様に、これら5シーンのSARデータを用いて、直播栽培と移植栽培の圃場でそれぞれ分割し、栽培形態ごとのモデルを作成した。その結果、いずれのモデルも全データモデルより精度は高く、特に移植栽培のモデルでは決定係数が0.5を上回っており、高い精度が得られた。この結果から、衛星進行方向だけでなく栽培形態も考慮に入れることにより、モデルの精度の向上が期待できることが示された。

以上の結果から、SARデータから水稻の穂含水率を推定するためには、衛星進行方向をアセンディング軌道に固定したうえでモデルを作成すると、良好な精度を得られることが示唆された。また、観測時期の都合によりディセンディング軌道のデータを使用するとしても、直播栽培と移植栽培の区別を行ってモデルを作成することにより、良好な精度で穂含水率を推定できる可能性も示唆された。なお、ほぼすべてのモデルでエントロピが変数として選択されていたため、特に水稻の穂含水率との間に関係性が深い変数であることが示され

た。

つづいて、SAR データと光学センサデータを併用した推定モデルの作成を行った。2017 年の 1 シーンの光学センサデータを使用して、NDVI から推定モデルを作成したところ、高い決定係数が得られなかった。しかし、1 点の外れ値を除くことで精度の向上がみられた。この 1 圃場は他の圃場と異なり、湛水されていたことから、光学センサのみを使用したモデルにおいては湛水の有無が精度を左右する可能性が示唆された。

そこで、SAR データと光学センサ 1 データを併用してモデルの作成を行ったところ、決定係数は 0.9 を超え、大幅な精度の向上がみられた。この結果から、両者を併用することで、湛水の有無の影響を減らすことができ、精度の向上に効果があると考えられる。したがって、水稲の穂含水率を推定するためには、SAR データと光学センサデータを併用することが最も良好な方法であると結論づけた。しかし、本地域は夏季に曇天となることが多く、対象とした 4 年間でも 1 シーンしか光学センサデータを取得できなかったことから、先述の SAR データのみによるモデルを有効に活用することが期待される。また、本章でも光学センサデータが少なかったことから、他の時期にも同様の精度が得られるかは不明であり、今後もデータを取得してモデルの改良に努めることも必要となる。

最後に、2016 年から 2018 年までの 5 シーンの SAR データから作成した 5 つのモデルについて、2019 年の 2 シーンの SAR データを用いた精度の検証を行った。過去年のデータを用いて作成したモデルを翌年のデータで検証することにより、今後このモデルを使用していくことが可能であるか検討することができる。その結果、全データモデルについては RMSE が 6%程度となり、良好な精度を得ることができた。全データモデルは作成した際の決定係数は低かったが、様々な時期や条件の情報が含まれていたことから、別の年のデータにも良好な適合がなされたと考えられる。

条件ごとに分割したモデルの検証では、直播栽培で RMSE が 4%以下の良好な結果が得られた。一方、他の 3 つのモデルは全て RMSE が 10%を上回り、全データを用いたモデルと比較して良好な結果を得ることができなかった。その理由として、ディセンディング軌道によるモデルは作成した時点で精度が低かったことから、検証においても高い精度が得られなかったと考えられる。したがって、ディセンディングのような朝露が影響する可能性のあるデータに対応しにくい点が問題として挙げられる。また、2019 年のアセンディング軌道のデータに関しては、入射角がモデルを作成した際に使用したデータより約 10°大きいデー

タであったことが要因として考えられる。したがって、衛星進行方向や栽培形態だけでなく、入射角についてもモデルの情報に取り入れる必要があると考えられる。移植栽培と直播栽培のモデルで検証の精度が大きく異なっていたことから、特に移植栽培の圃場ではこれら入射角や朝露などの影響を受けやすい可能性が示唆される。

以上の結果から、SAR と光学センサを併用したモデルは最も高い精度が得られる可能性がある一方で、光学センサが観測できない場合に SAR のみでもある程度良好なモデルを作成できることが示された。とくに直播栽培の圃場のみを対象としたモデルについては汎用性が高く、有効であった。移植栽培の圃場においても、入射角が  $30^\circ$  程度のアセンディング軌道で観測されたモデルであれば、ある程度の精度が得られる可能性が示された。

## 7 Cバンド SAR およびダミー変数を使用した水稻穂含水率の推定

### 7.1 背景と目的

前章において、RADARSAT-2により観測されたCバンド4偏波のSARデータを用いて、水稻の穂含水率を推定する重回帰モデルを作成したところ、1つのモデルでは高い精度が得られず、高い精度のモデルを作成するためには栽培形態や衛星進行方向でモデルを分ける必要があった。また、含水率の推定を行うにあたっては、衛星進行方向や入射角、栽培形態などの様々な要因が影響していると考えられるため、今後さらに高い精度のモデルを作成するためには、これらの要因のパターンごとに条件を組み合わせ、モデルを細分化して作成しなければならない可能性がある。しかしながら、複数のモデルを作成して選択する労力や、使用時におけるヒューマンエラーの発生も懸念され、水稻の穂含水率を推定するモデルを実利用する観点から、細分化したモデルは現実的な方策でないと考えられる。

そこで、衛星進行方向や入射角、栽培形態など、衛星データや圃場ごとの条件の差異を勘案するため、ダミー変数を使用したモデルの作成を行い、より少ないモデルの数で簡潔に水稻の穂含水率を推定する手法について検討した。ダミー変数とは、数値ではないデータを数値に置き換えた変数のことである。本章では、衛星データの入射角、衛星進行方向、栽培形態のそれぞれについて、データを0と1の数値として表し、説明変数として使用した。このダミー変数3変数と、6章と同様にSARデータから算出された9変数をもとに、重回帰分析より穂含水率推定モデルの作成と検証を行った。

本章ではまず、全てのSARデータのみによるモデルの作成を行い、この結果と他の手法によるモデルの作成結果について比較した。他の手法で作成したモデルは、前章と同様に、衛星進行方向を考慮したモデル2種類と、栽培形態を考慮したモデル2種類である。また、以上の4種類のモデルおよび全データを使用したモデルに、入射角、栽培形態、衛星進行方向の3種類のダミー変数をそれぞれ加え、再度モデルを作成した。これら9種類のモデルの精度を算出し、全てのSARデータのみで作成したモデルと比較を行い、ダミー変数を使用することの有効性について検討した。最後に、最も有効であったモデルについて5分割による交差検証を行い、RMSEを算出した。この結果から、モデルの汎化性能について考察し、このモデルの実用可能性について判断した。

## 7.2 研究方法

### 7.2.1 使用データおよび処理方法

本章の対象地域および対象作物，対象圃場は6章と同じである。使用した SAR 衛星データは，Cバンドのセンサを搭載した RADARSAT-2 の4偏波データである。SAR データには，6章と同様に Fine Quad モードにより観測されたものを使用した。ピクセルスペーシングは8.0m，処理レベルは Single Look Complex (SLC) である。表 7-1 にそれぞれの衛星データの観測時期，入射角，衛星進行方向と，現地調査の日程を示す。本章でも，入射角が約 30°の FQ9，FQ11 と，入射角が約 40°の FQ19 を使用した。また，FQ9 と FQ19 はアセンディング軌道によって観測されており，FQ11 はディセンディング軌道によって観測されている。それぞれの衛星データは出穂期から成熟期の間で観測されたものを使用したため，各年とも8月下旬から9月上旬までのものとなっている。

推定モデルを作成するにあたり，RADARSAT-2 データの前処理には，前章までと同様にノイズ除去，マルチルック処理，オルソ補正を行った。このように調整した SAR データを用いて，モデルに使用する変数を算出した。6章と同様に HH，HV，VV 偏波の後方散乱係数と，Van Zyl *et al.* (2011) による散乱モデル分解法の散乱成分の2回散乱 (Pd)，体積散乱 (Pv)，表面散乱 (Ps)，Cloude-Pottier による固有値解析法の散乱成分 (Cloude and Pottier, 1996) のエントロピ (H)，アニソトロピ (A)，アルファ角 ( $\alpha$ ) の計9変数を算出した。

SAR による9変数は，6章と同様に圃場単位で穂含水率を把握するため，各年の調査圃場

表 7-1 使用した SAR データおよび現地調査日程

年	日付	入射角	衛星進行方向	現地調査日
2016年	09月03日	FQ11 (30.2~32.0° )	ディセンディング	09月04日
2017年	08月26日	FQ9 (28.0~29.8° )	アセンディング	08月26日
	08月29日	FQ11 (30.2~32.0° )	ディセンディング	08月29日
2018年	08月21日	FQ9 (28.0~29.8° )	アセンディング	08月24日
	08月24日	FQ11 (30.2~32.0° )	ディセンディング	
2019年	08月20日	FQ11 (30.2~32.0° )	ディセンディング	08月20日
	08月26日	FQ19 (38.3~39.8° )	アセンディング	08月27日

の境界を示す GIS データを使用し、バッファの除去や平均値の抽出を行った。

作成するモデルの説明変数として、SAR より算出した以上の 9 変数のほか、3 種類のダミー変数を使用した。ダミー変数とは、数値ではないデータを数値に置き換えた変数のことである。本章では、衛星データの入射角 (FQ)、衛星進行方向 (AD)、栽培形態 (S) のそれぞれについて、データを 0 と 1 の数値として表し、説明変数として使用した。これらのダミー変数について、表 7-2 に示す。衛星データの入射角については数字で表すことができるが、圃場によって入射角は僅かに異なることから、本章では便宜上入射角が約 30°と小さい FQ9 および FQ11 を 0、入射角が約 40°と大きい FQ19 を 0 として使用している。

## 7.2.2 解析方法

本章でもモデルの作成には重回帰分析を使用した。目的変数は水稻の穂含水率 (MCE) であり、こちらも 6 章と同様の日程で現地調査を行い、同様の手法で算出した。説明変数にステップワイズ法の一つである変数減少法を用いて、説明変数の選択を行った。ステップワイズ法の評価指標には赤池情報量基準 (AIC) を使用した。次に、選択されたすべての変数の VIF 統計量を算出し、VIF 統計量が 5 を超えている変数は除き、残りの変数を用いてモデルを作成した。

作成したモデルの精度の評価には、自由度調整済み決定係数 ( $R^2$ ) および RMSE をそれぞれ使用した。ステップワイズ法によって選択された変数については、標準偏回帰係数を算出して精度を評価した。モデルの検証における精度評価には、RMSE のみを使用した。モデルの作成および検証には、フリー統計ソフトである R (R Core Team, 2016) を使用した。

表 7-2 使用したダミー変数

変数	ダミー変数の値	
	0	1
入射角 FQ	FQ9/11	FQ19
衛星進行方向 AD	アセンディング	ディセンディング
栽培体系 S	直播	移植

### 7.2.3 解析の流れ

本章ではまず、5種類の SAR データによる推定モデルと、5種類の SAR データおよびダミー変数による推定モデルを作成し、精度の評価や比較を行った。次に、最も有効であると判断したモデルについて、交差検証を行い、汎化性能の評価を行った。最後に、このモデルを使用して水稻の穂含水率の予測値を示す推定マップを作成した。詳細には、以下の 1)~3) で示す 3 処理で解析を進めた。

#### 1) SAR 変数とダミー変数を使用した水稻穂含水率の推定モデルの作成

まず、SAR データより算出された 9 変数のみを用いて、すべての圃場を対象とした水稻の穂含水率の推定モデルを作成した。本章では、前章で作成した 2016 年から 2018 年までのデータに 2019 年のデータを加え、全 7 シーンのデータでモデルを作成した。この結果を基準として、様々な条件で作成したモデルの精度を評価した。

6 章において、衛星進行方向でモデルを分けることにより、精度の向上がみられた。そこで、2016 年から 2019 年までの SAR データのうち、アセンディング軌道で観測された FQ9 データおよび FQ19 データ (計 3 シーン) またはディセンディング軌道で観測された FQ11 データ (計 4 シーン) のいずれかを使用し、衛星進行方向ごとに推定モデルを作成した。各衛星進行方向について、SAR による 9 変数のみを使用したモデルと、ダミー変数を加えたモデルのそれぞれを作成した。アセンディング軌道によるモデルでは入射角 (FQ) と栽培形態 (S) のダミー変数 2 変数を使用し、ディセンディング軌道によるモデルでは栽培形態 (S) のダミー変数 1 変数のみを使用した。以上の方法で作成した 4 種類のモデルについて、すべてのデータを使用したモデルとの精度比較や、ダミー変数を使用することの有効性について検討した。

また、6 章においては、栽培形態でモデルを分けることによっても精度の向上がみられた。そこで、2016 年から 2019 年までの全 7 シーンの SAR データを使用し、直播栽培の圃場のみを使用したモデルと、移植栽培の圃場のみを使用したモデルをそれぞれ作成した。各栽培形態のモデルについて、SAR による 9 変数のみを使用したモデルと、ダミー変数を加えたモデルのそれぞれを作成した。ダミー変数には、入射角 (FQ) と衛星進行方向 (AD) の 2 変数を使用した。以上の方法で作成した 4 種類のモデルについて、すべてのデータを使用したモデルとの精度比較や、ダミー変数を使用することの有効性について検討した。

最後に、2016 年から 2019 年までの全 7 シーンの SAR データから算出された 9 変数と、

全3種類のダミー変数のすべてを使用してモデルを作成した。作成したモデルの精度から、ダミー変数を使用することで1つのモデルで水稻の穂含水率を推定することができるかについて検討した。

## 2) SAR変数とダミー変数を使用した水稻穂含水率の推定モデルの検証

1) で作成したモデルのうち、最も有効であると判断したモデルについて交差検証を行った。交差検証には、先行研究 (Ndikumana *et al.*, 2018) と同様に、5分割による検証手法を用いた。まず、すべての圃場のデータを5分割し、5つのグループに分けた。このうち4つのグループに含まれるデータをすべて用いてモデルを作成し、残りの1グループに含まれるデータを用いて検証を行った。これを5回すべてのパターンで繰り返し、RMSEを平均した。モデル作成に当たっては、1) で作成された当該のモデルにおいて選択された変数のみを使用したモデルと、すべての変数で再度ステップワイズをし直したモデル、当該のモデルにダミー変数を使用していないモデルの計3種類でそれぞれ交差検証を行った。以上の方法を元に、このモデルの汎化性能の有無について判断し、有効なモデルであるか判断した。

## 3) SAR変数とダミー変数を使用した水稻穂含水率の推定マップの作成

1) で作成したモデルのうち最も有効であると判断したモデルを、水稻圃場のポリゴンに適用し、水稻の穂含水率の予測値を示す推定マップを作成した。この圃場ポリゴンは、2019年の2農業法人すべての水稻圃場の区画を示しており、品種の属性情報が含まれている。モデルを2019年の2シーンのSARデータに当てはめ、2シーンの推定マップを作成し、有効なマップが作成されているか検討した。

## 7.3 結果と考察

### 7.3.1 SAR変数とダミー変数を使用した水稻穂含水率の推定モデル作成

#### 1) 全てのSARデータを用いたモデルの作成

2016年から2019年までの4年間に観測されたすべてのSARデータを用いたモデルを作成し、実測値と推定値を比較した結果を図7-1に示す。決定係数は0.148、RMSEは8.07%となっており、高い精度を得ることができなかった。6章において、2016年から2018年の3年間のデータでも同様のモデルを作成しているが、決定係数はほぼ同じ値となっており、

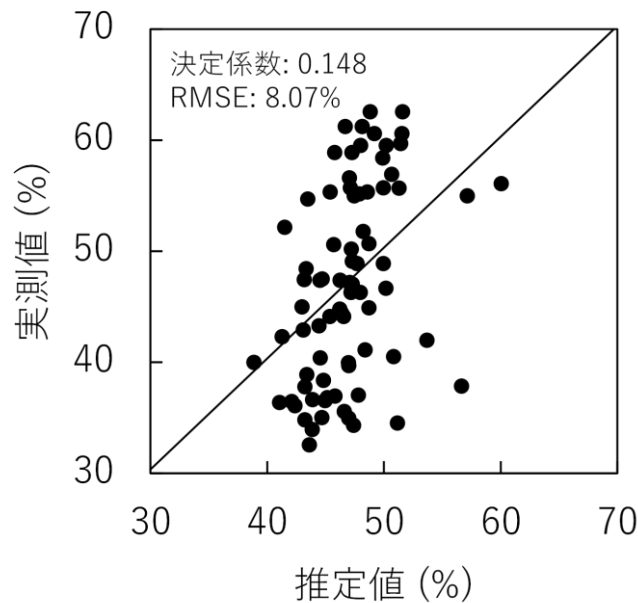


図 7-1 全データを使用したモデルによる穂含水率の推定値と実測値の関係

次年度のデータを追加しても精度の向上や低下はみられなかった。また、6章と同様に、実測値が高い圃場では推定値が低く、過小評価されている傾向がみられ、実測値が低い圃場では推定値が高く、過大評価されている傾向がみられた。以上の結果より、4年間のデータを用いた結果においても、穂含水率を推定するためにすべてのデータを用いた高い精度のモデルを作成することは困難であった。

作成されたモデルを式 (7-1) に示す。

$$\text{MCE} = 1.895 \times \text{VV} - 33.295 \times \text{H} + 94.530 \quad (7-1)$$

モデルには、6章で示した3年間のデータで作成したモデルと同様に、VV 偏波後方散乱係数とエントロピの2変数が選択された。この2変数について、目的変数の推定に寄与している変数を標準偏回帰係数で比較した(表 7-3)。その結果、VV 偏波後方散乱係数は標準偏回帰係数が0.2を上回っていたほか、エントロピは-0.2を下回っており、両者が同程度の影響の強さであった。前章までの結果でも示されているように、VV 偏波はHH 偏波より作物体に関する情報の収集に適しているとされており (Englhart *et al.*, 2011), 本研究でも同様に

表 7-3 全データを使用したモデルに選択された各変数の標準偏回帰係数

変数	標準偏回帰係数
VV	0.293
H	-0.265

推定モデルへの寄与が大きかったといえる。エントロピは散乱のランダム性を表す変数であるため、穂含水率の多寡によって変化したと考えられる。

## 2) 衛星進行方向別のモデルの作成

4年間の SAR データのうち、ダミー変数を使用せずにディセンディング軌道で観測された FQ11 データ (計 4 シーン) のみを使用して作成したモデルと、アセンディング軌道で観測された FQ9 データおよび FQ19 データ (計 3 シーン) のみを使用して作成したモデルについて、実測値と推定値を比較した各結果を図 7-2 に示す。ディセンディング軌道によるモデルとアセンディング軌道によるモデルの決定係数はそれぞれ 0.145, 0.389 となった。本章でも衛星進行方向でモデルを分けることにより、アセンディング軌道によるモデルは全データを使用したモデルと比較して高い精度となったが、6章で作成した3年間のデータによるモデルより大幅に精度の低下がみられた。この原因として、2019年8月26日のデータの

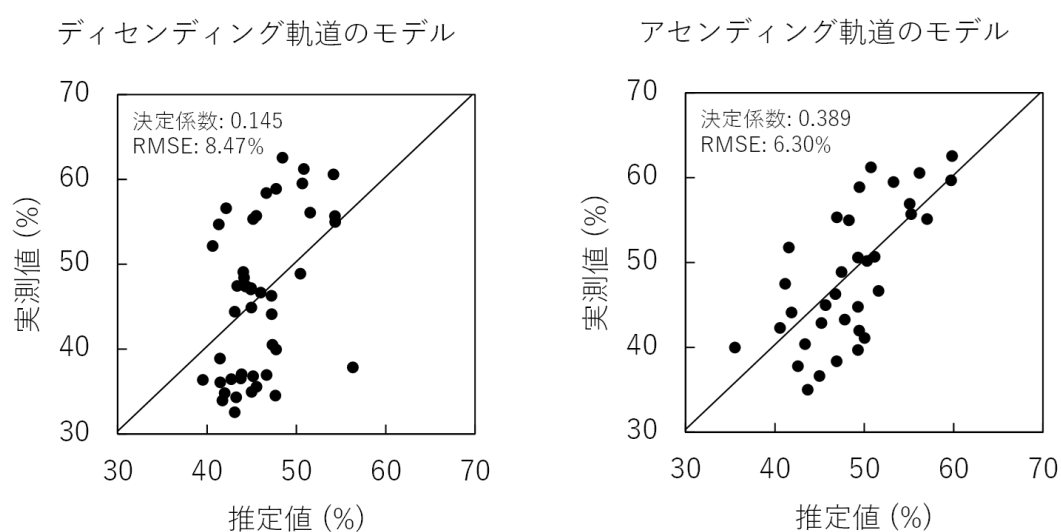


図 7-2 衛星進行方向ごとに作成したモデルによる穂含水率の推定値と実測値の関係

み、約 10°入射角の大きい FQ19 であったことが挙げられる。入射角が異なったことで散乱の形態が異なり、精度の低下に繋がった可能性が考えられる。6 章でも示したように、FQ9 で作成したモデルで、FQ19 のデータの検証は良好な結果でなかった。したがって、衛星進行方向だけでなく衛星の入射角についても検討していく必要があり、両者を併用したモデルの構築を SAR 変数のみで行うことは困難であるといえ、ダミー変数の必要性が示唆される。また、ディセンディング軌道によるモデルは全データを使用したモデルとほぼ決定係数や RMSE に差異はみられなかった。6 章でもディセンディング軌道はアセンディング軌道と比較して低い精度となっていたが、この理由はデータ観測時間が異なっていることが考えられる。本地域では、ディセンディング軌道の観測は早朝、アセンディング軌道の観測は夕刻に行われているため、ディセンディング軌道の観測時間のみ朝露が発生しており、レーダの散乱に影響を与え、精度の低下を及ぼしていることが考えられる。以上の結果より、衛星進行方向ごとに作成した穂含水率の推定モデルに関しては、特にアセンディング軌道によるモデルは全データを使用したモデルより高い決定係数は得られたものの、高い精度のモデルを作成することは困難であった。

ディセンディング軌道によるモデルおよびアセンディング軌道によるモデルを、それぞれ式 (7-2)、式 (7-3) に示す。

$$\text{MCE} = 1.751 \times \text{VV} - 0.518 \times \text{Alp} + 89.719 \quad (7-2)$$

$$\begin{aligned} \text{MCE} = & 0.021 \times \text{HV} + 0.617 \times \text{Pd} + 1.244 \times \text{Ps} - 59.065 \times \text{H} \\ & + 0.451 \times \text{Alp} - 43.425 \end{aligned} \quad (7-3)$$

ディセンディング軌道によるモデルには、6 章の 3 年間のデータによるディセンディング軌道によるモデルと同様にアルファ角が選択されたほか、新たに VV 偏波後方散乱係数も選択され、計 2 変数となった。アセンディング軌道によるモデルも同様に 2 回散乱、エントロピが選択されたほか、新たに HV 偏波後方散乱係数や表面散乱、アルファ角も選択され、計 5 変数となった。これらの変数について、目的変数の推定に寄与している変数を標準偏回帰係数で比較した (表 7-4, 7-5)。ディセンディング軌道によるモデルについては、特に VV 偏波後方散乱係数は標準偏回帰係数が 0.2 を上回っていたほか、アルファ角も標準偏回帰係数

表 7-4 ディセンディング軌道のデータのみを使用したモデルに選択された

各変数の標準偏回帰係数

変数	標準偏回帰係数
VV	0.264
Alp	-0.243

表 7-5 アセンディング軌道のデータのみを使用したモデルに選択された

各変数の標準偏回帰係数

変数	標準偏回帰係数
HV	0.004
Pd	0.222
Ps	0.354
H	-0.493
Alp	0.269

が-0.2を下回っていた。アセンディング軌道によるモデルについては、エントロピの標準偏回帰係数が-0.4を下回っており、最も寄与が大きかった。ディセンディング軌道によるモデルについては、6章の3年間によるディセンディング軌道によるモデルと選択された変数や寄与の大きい変数の傾向が近かった。一方、アセンディング軌道によるモデルについては、6章のモデルで高い標準偏回帰係数を示していたHV 偏波後方散乱係数で標準偏回帰係数が低く、体積散乱は選択されなかった。アセンディング軌道によるモデルは、先述のとおり異なる入射角のデータを含めていることから、これらの変数が入射角によって値が異なるため、モデルの変数に選択されなかったと考えられる。

以上のように、SARの9変数のみによる衛星進行方向別のモデルでは良好な結果を得られなかったことから、入射角および栽培形態を示すダミー変数を使用し、計10変数または11変数でモデルを再構築した。アセンディング軌道のモデルでは入射角(FQ)と栽培形態(S)のダミー変数2変数を追加した。ディセンディング軌道のモデルでは、すべて入射角の等しいFQ11データのみを用いているため、栽培形態(S)のダミー変数1変数のみを追加した。このようにして作成したモデルによる推定値と実測値を比較した各結果を図7-3に示

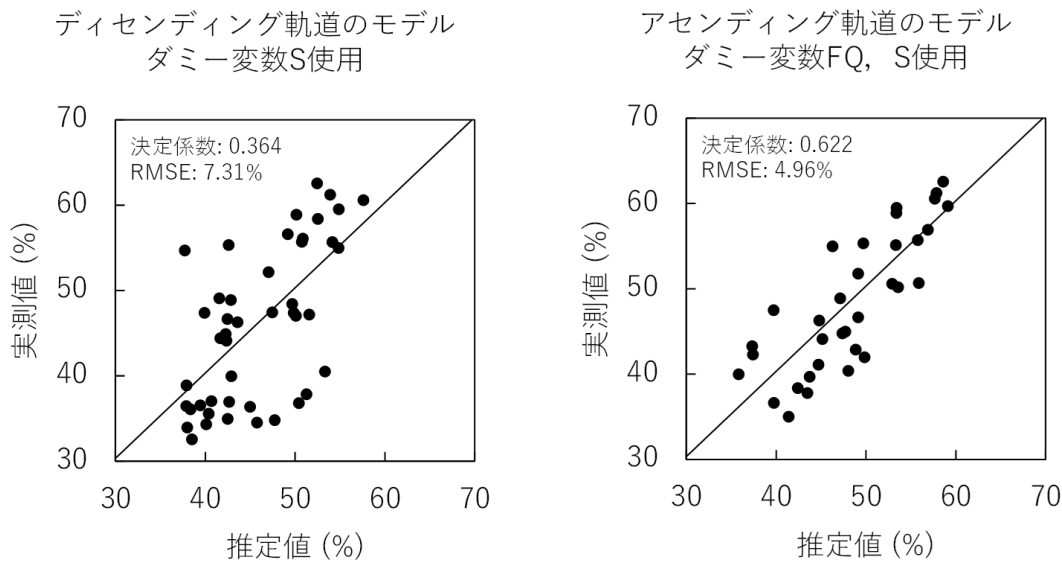


図 7-3 ダミー変数を使用し衛星進行方向ごとに作成したモデルによる  
穂含水率の推定値と実測値の関係

す。ディセンディング軌道によるモデルとアセンディング軌道によるモデルの決定係数はそれぞれ 0.364, 0.622 となり、ダミー変数を加えることによって衛星進行方向ごとのモデルにおける精度が向上した。とくにアセンディング軌道によるモデルは決定係数が 0.6 を超えており、より精度が向上したが、6 章で得られた決定係数と同等の結果となった。この原因として、先述のモデルと同様に、散乱の形態が異なる別の入射角のデータを含めたことが、精度の低下に繋がった可能性が考えられる。また、ディセンディング軌道によるモデルについては、6 章の 3 年間によるモデルおよび先述の 4 年間によるモデルにおいて、実測値が高い圃場では推定値が低く、過小評価されている傾向がみられ、実測値が低い圃場では推定値が高く、過大評価されている傾向がみられた。一方、ダミー変数を加えた本モデルでは、これまでのモデルと比較してこの傾向が緩和されており、より有効なモデルが作成されたといえる。以上の結果より、衛星進行方向別のモデルにおいては、ダミー変数の使用が有効であると考えられる。

ダミー変数を使用したディセンディング軌道によるモデルおよびアセンディング軌道によるモデルを、それぞれ式 (7-4)、式 (7-5) に示す。

$$\text{MCE} = 1.283 \times \text{Ps} - 9.211 \times \text{S} - 29.845 \quad (7-4)$$

$$\begin{aligned} \text{MCE} = & 2.781 \times \text{HV} - 38.644 \times \text{H} + 34.544 \times \text{A} - 9.706 \times \text{S} \\ & + 118.223 \end{aligned} \quad (7-5)$$

ディセンディング軌道によるモデルには、表面散乱と栽培形態を示すダミー変数が選択され、計2変数となった。アセンディング軌道によるモデルにはHV 偏波後方散乱係数、エントロピ、アニソトロピの3変数のほか、栽培形態を示すダミー変数が選択された。これらの変数について、目的変数の推定に寄与している変数を標準偏回帰係数で比較した(表7-6, 7-7)。ディセンディング軌道によるモデルについては、栽培形態のダミー変数は標準偏回帰係数が-0.5を下回っていたほか、表面散乱は標準偏回帰係数が0.4を上回っていた。今回のモデルで選択された表面散乱は、主に地表面の影響で値が変化する。2019年の早い時期のデータなどを加えたことにより、湛水されている圃場など土壌条件や生育状態が異なるデータがモデルに含まれるようになった。そのため、このような生育状態の異なる圃場の情報を補完するために有効であると考えられる。しかし、表面散乱以外のSAR変数は選択されなかったことから、ディセンディング軌道のデータのみを使用したモデルでは、HV 偏波後方散乱係数や体積散乱、エントロピなどの生育に関する変数が朝露などの影響を受け、水稲

表 7-6 ディセンディング軌道のデータのみを使用しダミー変数を加えたモデルに  
選択された各変数の標準偏回帰係数

変数	標準偏回帰係数
Ps	0.414
S	-0.502

表 7-7 アセンディング軌道のデータのみを使用しダミー変数を加えたモデルに  
選択された各変数の標準偏回帰係数

変数	標準偏回帰係数
HV	0.531
H	-0.323
A	0.207
S	-0.610

の穂含水率を良好に説明できなかつたと考えられる。

アセンディング軌道によるモデルについては、HV 偏波後方散乱係数で標準偏回帰係数が 0.5 を上回っていたほか、栽培形態のダミー変数は標準偏回帰係数が -0.6 を下回っていた。HV 偏波後方散乱係数のように作物体の大きさによって値が変化する変数は、特に穂含水率の推定に有効な変数であったと考えられる。HV 偏波後方散乱係数が有効であったという結果は、6 章でも同様にみられた。またエントロピに関しても、アセンディング軌道によるモデルで選択されていただけでなく、6 章においてもほぼすべてのモデルで選択されていたことから、重要な変数であるといえる。エントロピを用いることにより、穂の水分の差異によって散乱形態が変化する様相を補足することができたと考えられる。また、双方のモデルでは栽培形態を示すダミー変数が選択され、標準偏回帰係数の絶対値も高い値を示していた。したがって、穂含水率を推定するモデルを衛星進行方向別に作成するにあたっては、SAR の変数だけでなく栽培形態をダミー精度として説明変数に加えることが、精度を向上させる重要な方策となりうると考えられる。

### 3) 栽培形態別でのモデルの作成

4 年間の全 7 シーンの SAR データを使用し、直播栽培、移植栽培のみをそれぞれ対象としてダミー変数を使用せずに水稻の穂含水率の推定モデルを作成した。作成したモデルについて、実測値と推定値を比較した各結果を図 7-4 に示す。直播栽培によるモデルと移植栽培によるモデルの決定係数はそれぞれ 0.391, 0.203 となった。直播栽培は移植栽培と比較して高い精度が得られた。この理由として、直播栽培は移植栽培より茎の生育密度が大きくなっている圃場が多く、条間の影響が移植栽培より少ないことが考えられる。移植栽培では条を横切る方向に茎や穂の密度が変化し、圃場ごとに方向が多様であるため、マイクロ波の散乱に条方向による異なる影響が生じた可能性があるのに対し、直播栽培では空間分布が圃場内でほぼ一様であるため方向の影響が小さかった、もしくはなかったことが考えられる。6 章において、各栽培形態で作成したモデルは全てのデータを使用したモデルより高い精度が得られていたが、本結果では移植栽培においてほぼ精度の向上はみられなかった。この原因として、2019 年 8 月 20 日の早い時期のデータを加えた点や、2019 年 8 月 26 日の入射角が異なるデータを加えた点が挙げられる。生育密度が大きい直播栽培については、これらのデータを加えても精度の大幅な低下はみられなかったが、条方向の影響が大きい移植栽培については、6 章で使用したデータと時期や入射角が異なるデータを加えたことによって、

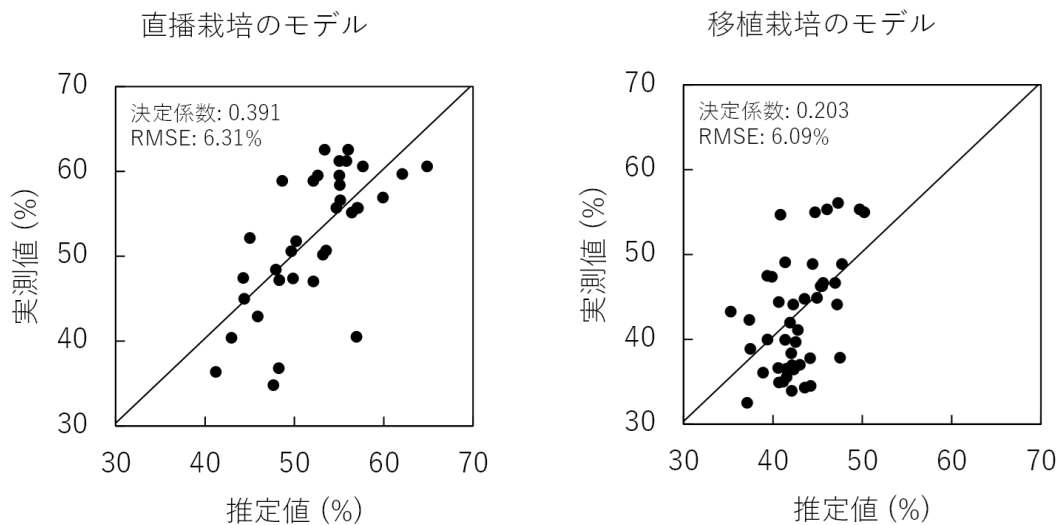


図 7-4 栽培形態ごとに作成したモデルによる穂含水率の推定値と実測値の関係

更に大きく影響を受けたと考えられる。また、対象とした移植栽培の圃場では4種類の品種が栽培されているが、これらの品種間でも散乱の傾向が異なる可能性がある。したがって、栽培形態別のモデルでも SAR 変数のみで穂含水率を推定することが困難であるといえ、特に移植栽培では衛星進行方向や入射角などの影響を含めるためにダミー変数の必要性が大きいといえる。

直播栽培によるモデルおよび移植栽培によるモデルを、それぞれ式 (7-6)、式 (7-7) に示す。

$$\begin{aligned} \text{MCE} = & 2.408 \times P_v - 56.142 \times H + 37.563 \times A - 0.636 \times A_{lp} \\ & - 55.390 \end{aligned} \quad (7-6)$$

$$\text{MCE} = 2.357 \times H_V + 41.662 \times A + 68.666 \quad (7-7)$$

直播栽培のモデルは、6章の3年間のデータで作成した直播栽培のモデルと同様にエントロピが選択されたほか、新たに体積散乱、アニソトロピ、アルファ角の3変数が選択され、計4変数となった。移植栽培のモデルも6章の3年間のデータで作成した移植栽培のモデルと同様に HV 偏波後方散乱係数が同様に選択されたほか、新たにアニソトロピも選択され、計

2変数となった。これらの変数について、目的変数の推定に寄与している変数を標準偏回帰係数で比較した(表 7-8, 7-9)。直播栽培については、体積散乱で標準偏回帰係数が 0.3 を上回ったほか、エントロピとアルファ角でも-0.3 を下回った。移植栽培については、HV 偏波後方散乱係数で標準偏回帰係数が 0.4 を上回っていた。直播栽培でエントロピの寄与が大きい結果となったのは、6章で作成した3年間のデータによるモデルと同様の結果であり、本章におけるこれまでの結果とも同様であった。移植栽培に関しても、6章で作成した3年間のデータによるモデルと同様に HV 偏波後方散乱係数が良好な変数となり、水稻自体の状態を得るために有用な変数として寄与していたことが示唆された。

以上のことから、SAR の 9 変数のみによる栽培形態別のモデルでは、特に移植栽培モデルにおいて良好な結果を得られなかった。そこで、入射角 (FQ) および衛星進行方向 (AD) を示すダミー変数を追加し、計 11 変数でモデルを再構築した。作成したモデルについて、実測値と推定値を比較した各結果を図 7-5 に示す。直播栽培によるモデルと移植栽培によるモデルの決定係数はそれぞれ 0.362, 0.235 となった。移植栽培によるモデルでは、ダミー変数を使用したことによって精度の向上がわずかにみられたが、大幅な精度の向上は達成されなかった。また、直播栽培によるモデルはダミー変数を使用したことによって精度の低下がみられた。したがって、栽培形態別のモデルでも SAR 変数にダミー変数を加えることにより、衛星進行方向や入射角などの影響を低減させることが可能となるが、特にダミー変数

表 7-8 直播栽培のみを使用したモデルに選択された各変数の標準偏回帰係数

変数	標準偏回帰係数
Pv	0.385
H	-0.332
A	0.278
Alp	-0.330

表 7-9 移植栽培のみを使用したモデルに選択された各変数の標準偏回帰係数

変数	標準偏回帰係数
HV	0.461
A	0.334

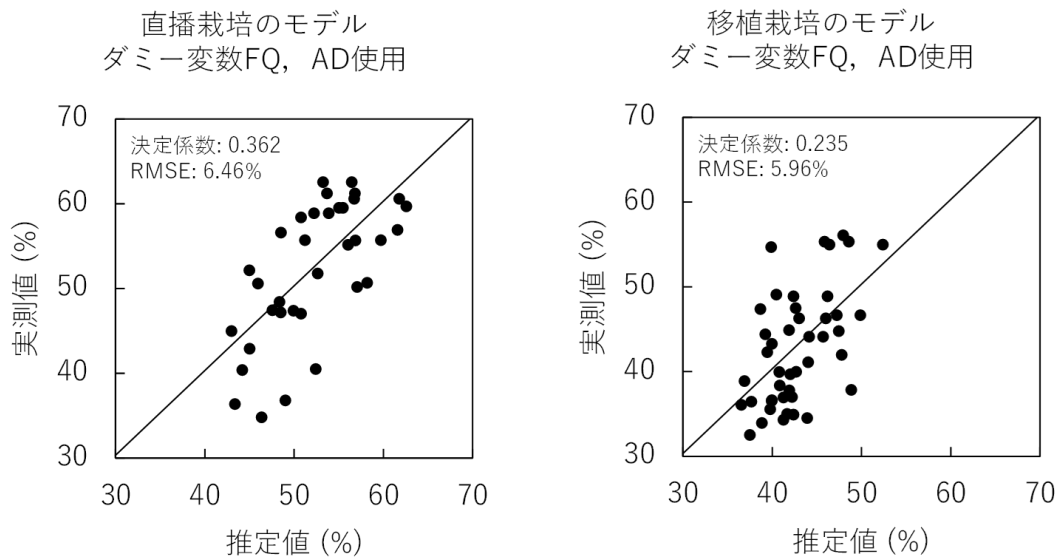


図 7-5 ダミー変数を使用し栽培形態ごとに作成したモデルによる穂含水率の推定値と実測値の関係

を用いなくてもある程度良好な精度が得られていた直播栽培では、ダミー変数を用いる必要が低いと考えられる。また、6章では移植栽培のモデルの決定係数が0.52と比較的高い精度であったことから、移植栽培のモデルにおける穂含水率の推定には、使用する衛星データの時期を吟味する必要があると考えられる。

ダミー変数を使用した直播栽培によるモデルおよび移植栽培によるモデルを、それぞれ式(7-8)、式(7-9)に示す。

$$\begin{aligned} \text{MCE} = & 1.024 \times \text{HH} + 1.233 \times \text{HV} - 0.674 \times \text{Alp} - 7.258 \times \text{FQ} \\ & - 5.807 \times \text{AD} + 109.798 \end{aligned} \quad (7-8)$$

$$\begin{aligned} \text{MCE} = & 2.338 \times \text{HV} + 0.644 \times \text{Pd} + 0.537 \times \text{Ps} + 7.323 \times \text{FQ} \\ & + 9.997 \end{aligned} \quad (7-9)$$

直播栽培のモデルは先述のモデルと同様にアルファ角が選択されたほか、新たにHH偏波後方散乱係数、VV偏波後方散乱係数と、入射角および衛星進行方向を示すダミー変数が選択され、計5変数となった。移植栽培のモデルも先述のモデルと同様にHV偏波後方散乱係

数が選択されたほか、2回散乱、表面散乱と入射角を示すダミー変数が選択され、計4変数となった。これらの変数について、目的変数の推定に寄与している変数を標準偏回帰係数で比較した(表7-10, 7-11)。直播栽培については、先述のダミー変数を使用していないモデルと同様に、アルファ角が特に標準偏回帰係数の値が良好であり、-0.3を下回っていたほか、入射角と衛星進行方向のダミー変数も-0.3を下回った。移植栽培についても、先述のダミー変数を使用していないモデルと同様に、HV偏波後方散乱係数の標準偏回帰係数が良好であり、0.4を上回っていたほか、入射角のダミー変数も0.3を上回った。HV偏波後方散乱係数はこれまでのモデルでも選択された場合に高い標準偏回帰係数が得られており、本モデルでも同様の結果となった。また、衛星進行方向のダミー変数は選択されずに入射角のダミー変数が選択されたことから、移植栽培においては入射角の影響は衛星進行方向の影響よりも大きく、モデル作成を行うにあたって入射角を加味する必要があることが示された。以上の結果より、穂含水率を推定するモデルを栽培形態別に作成するにあたっては、SARの変数だけでなく、入射角の導入が精度を向上させるうえで重要であり、特に移植栽培で精度

表 7-10 直播栽培のみを使用しダミー変数を加えたモデルに選択された

各変数の標準偏回帰係数	
変数	標準偏回帰係数
HH	0.193
VV	0.200
Alp	-0.350
FQ	-0.319
AD	-0.363

表 7-11 移植栽培のみを使用しダミー変数を加えたモデルに選択された

各変数の標準偏回帰係数	
変数	標準偏回帰係数
HV	0.457
Pd	0.245
Ps	0.238
FQ	0.373

を向上させるための方策となると考えられる。

#### 4) 全データにダミー変数を加えたモデルの作成

4年間のすべての SAR データから算出された 9 変数と、入射角 (FQ)、衛星進行方向 (AD)、栽培形態 (S) のダミー変数の 3 変数を用いて、計 12 変数によるモデルを作成した。作成したモデルについて、実測値と推定値を比較した結果を図 7-6 に示す。決定係数は 0.452、RMSE は 6.48%と高い精度が得られ、1) で作成したダミー変数を使用していない全データによるモデルと比較して、精度の向上がみられただけでなく、2) で作成したダミー変数を用いたアセンディング軌道によるモデルに次ぐ精度となった。ダミー変数を使用していないモデルでは、実測値が高い圃場では推定値が低く、過小評価されている傾向がみられ、実測値が低い圃場では推定値が高く、過大評価されている傾向がみられていた。しかし、ダミー変数を加えることにより、このような影響が緩和された。また、この結果は、3) で作成した栽培形態ごとのダミー変数を用いたモデルと比較すると良好な決定係数や RMSE であった。したがって、栽培形態ごとにモデルを分けなくても、ダミー変数を用いることで同程度の精度が得られるため、穂含水率の推定モデルを 1 つのモデルにまとめられることが示された。

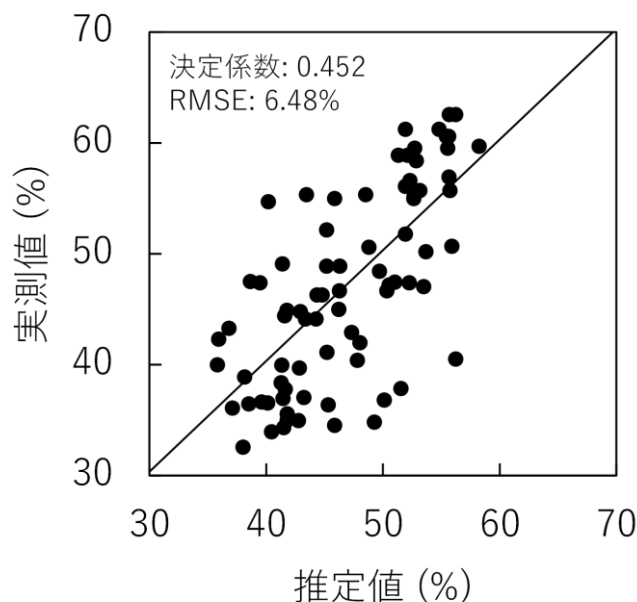


図 7-6 全データを使用しダミー変数を加えたモデルによる穂含水率の推定値と実測値の関係

作成されたモデルを式 (7-10) に示す。

$$\begin{aligned} \text{MCE} = & 1.728 \times \text{HV} + 1.197 \times \text{VV} - 31.149 \times \text{H} - 10.630 \times \text{S} \\ & + 121.485 \end{aligned} \quad (7-10)$$

モデルには 1) のモデルでも選択された VV 偏波後方散乱係数とエントロピの 2 変数が選択されたほか、HV 偏波後方散乱係数および栽培形態を示すダミー変数が選択され、計 4 変数となった。この 4 変数について、目的変数の推定に寄与している変数を標準偏回帰係数で比較した (表 7-12)。その結果、特に栽培形態を示すダミー変数は標準偏回帰係数が-0.6 を下回っており、特に穂含水率の推定に寄与しているといえる。ダミー変数では栽培形態のみが選択されたことから、入射角や衛星進行方向の情報よりも栽培形態の情報が本モデルでは重要となると考えられる。また、HV 偏波後方散乱係数は標準偏回帰係数が 0.2 を上回り、エントロピも-0.2 を下回った。このことから、これまでのモデルで最も多く選択されたエントロピや、選択された際に高い標準偏回帰係数となった HV 偏波後方散乱係数も穂含水率を推定するための有用な変数であることを本モデルでも示した。

以上の結果から、水稻の穂含水率を推定するモデルを作成するにあたり、①ダミー変数を使用していないモデルでは精度が低下する点、②栽培形態別のダミー変数を使用したモデルは全データのダミー変数を使用したモデルより精度が低い点、③衛星進行方向別のダミー変数を使用したモデルはディセンディング軌道によるモデルで精度が低い点の 3 点を勘案し、4) で作成した全データにダミー変数を加えたモデルが最も有効な水稻の穂含水率の推定モデルであると考えられる。

表 7-12 全データを使用しダミー変数を加えたモデルに選択された

各変数の標準偏回帰係数	
変数	標準偏回帰係数
HV	0.253
VV	0.185
H	-0.248
S	-0.608

### 7.3.2 SAR 変数とダミー変数を使用した水稻穂含水率の推定モデルの検証

7.3.1 で作成したモデルのうち、4) で作成した全データにダミー変数を加えたモデルが最も有効な水稻の穂含水率の推定モデルであると判断し、このモデルについて交差検証を行った。検証を行ったモデルは、4) で選択された4変数 (HV 偏波後方散乱係数, VV 偏波後方散乱係数, エントロピ, 栽培形態) からモデルを作成することの有効性を検討するために、この4変数のみを使用したモデル (ダミー変数あり4変数モデル) と、全12変数で再度ステップワイズ法を適用し直したモデル (ダミー変数あり再構築モデル)、1) で作成したダミー変数を使用しないモデルを作成した際に選択された2変数 (VV 偏波後方散乱係数, エントロピ) を使用したモデル (ダミー変数なしモデル) の3種類について、それぞれ交差検証を行った。以上の方法を元に、このモデルの汎化性能の有無について判断し、有効なモデルであるか判断した。

作成した各モデルにおける5回を平均した決定係数, RMSE と、検証結果における RMSE を、表 7-13 に示す。ダミー変数あり4変数モデルでは決定係数の平均が 0.451, RMSE の平均は 6.48%となり、4) で作成したモデルとほぼ同等の結果となった。ダミー変数あり再構築モデルについては、決定係数の平均が 0.459, RMSE の平均は 6.43%となり、ダミー変数あり4変数モデルよりわずかに高い精度となったものの、ほぼ変わらない結果となった。ダミー変数なしモデルは決定係数, RMSE と最も不良な結果となった。

検証結果については、ダミー変数ありの双方のモデルで比較したところ、ダミー変数あり4変数モデルの方がダミー変数あり再構築モデルより RMSE が 0.6%以上低く、平均は 6.70%となった。ダミー変数あり再構築モデルでは、5回のモデル作成において、アノトロピ, アルファ角, 衛星進行方向などの変数が選択されることがあった。ダミー変数あり再構築モデルで精度が低下した原因として、変数を変更したことにより決定係数や作成モデルの

表 7-13 全データを使用しダミー変数を加えたモデルの交差検証結果

モデル	作成結果		検証結果
	決定係数	RMSE (%)	RMSE (%)
ダミー変数なしモデル	0.145	8.09	8.09
ダミー変数あり4変数モデル	0.451	6.48	6.70
ダミー変数あり再構築モデル	0.459	6.43	7.36

RMSE は若干向上されたものの、モデルに過適合が生じてしまい、検証結果における RMSE が高くなったと考えられる。以上の結果から、検証でも大幅な精度の低下がみられず、水稻の穂含水率を推定するためのモデルとして有効であると判断できるだけでなく、4) で選択された HV 偏波後方散乱係数、VV 偏波後方散乱係数、エントロピ、栽培形態のダミー変数の計 4 変数で作成したモデルは、汎化性能が高いと考えられ、本研究における最も有用性の高いモデルであると判断した。

### 7.3.3 SAR 変数とダミー変数を使用した水稻穂含水率の推定マップの作成

7.3.2 で検証した結果より、4) で作成した全データにダミー変数を加えた 3 変数のモデルが最も有効な水稻の穂含水率の推定モデルであると判断し、このモデルをもとに地域全体の水稻の穂含水率の予測値を示す推定マップを作成した。地域全体の圃場の情報を得るため、水稻作付圃場のポリゴンを使用した。この圃場ポリゴンは、2019 年の 2 農業法人におけるすべての水稻作付圃場の区画を示している。また、このポリゴンには品種の属性情報が含まれているため、栽培形態を示すダミー変数を入力する際の参考とした。このポリゴンを使用して、2019 年の 2 シーンの SAR データから各 SAR 変数の平均値を算出した。以上の方法で算出した SAR 変数とダミー変数を、先述の 3 変数のモデルに当てはめ、2 シーンの推定マップをそれぞれ作成し、有効なマップが作成されているか検討した。

2019 年 8 月 20 日の穂含水率の推定マップと、2019 年 8 月 26 日の穂含水率の推定マップを図 7-7、図 7-8 に示す。これらの図より、各シーンのマップで含水率の高い圃場と低い圃場が一致している傾向がみられた。しかし、両マップを比較したところ、各圃場で含水率が上昇している傾向がみられた。8 月 20 日のマップは含水率が約 30~45%となっている圃場が多い一方で、8 月 26 日のマップでは含水率が 35~50%となっている圃場が多くみられた。これらのマップにおける含水率の予測値は、6 章で示した現地調査の時期の含水率と比較して、8 月 26 日の値はほぼ一致していたものの、8 月 20 日の値は現地調査の値より凡そ 10 ポイント低い値となっていた。8 月 20 日のマップで予測値が低下した原因として、特に移植栽培の圃場において、VV 偏波後方散乱係数の値が低い圃場が多くみられたことが考えられる。マップで対象とした地域において、HV 偏波後方散乱係数やエントロピの値は、経時変化による穂含水率の低下に伴いそれぞれ低下、上昇していた。これは、モデルにおける経時変化の挙動と同様である。一方、VV 偏波後方散乱係数は、経時変化による穂含水率の低下

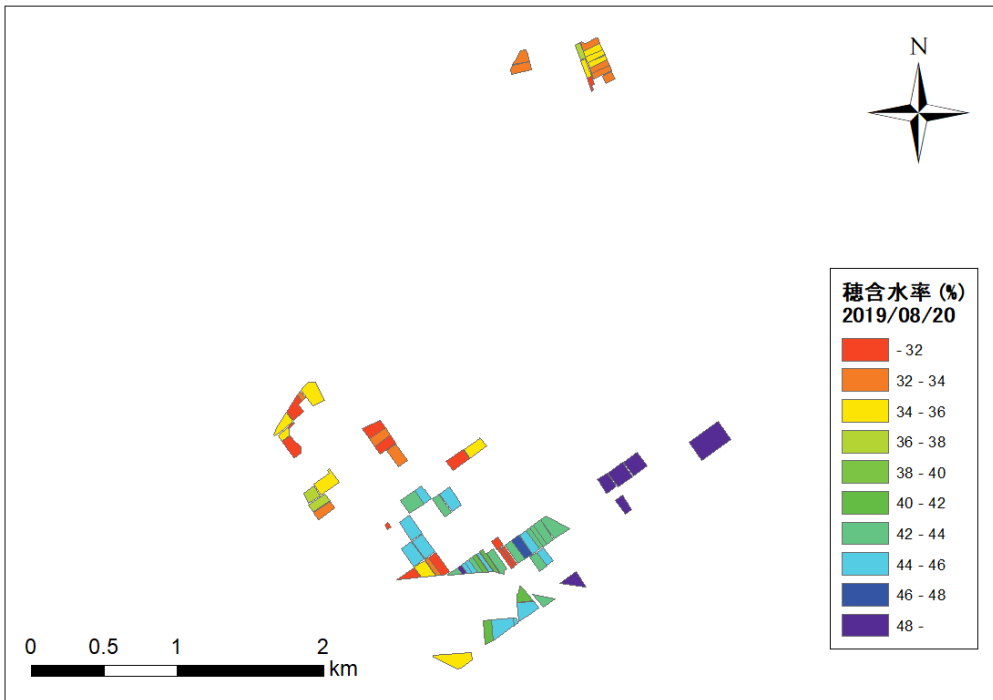


図 7-7 2019 年 8 月 20 日の穂含水率推定マップ

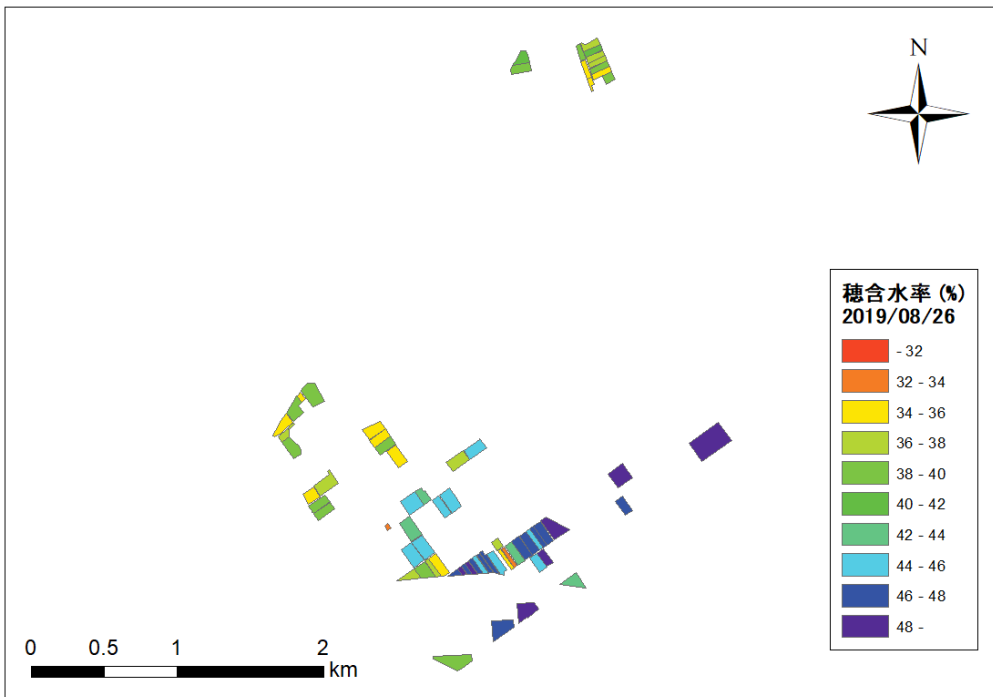


図 7-8 2019 年 8 月 26 日の穂含水率推定マップ

に伴いモデル内の説明変数でも低下するが、マップで対象とした圃場においては経時変化により上昇しており、逆の挙動を示していた。VV 偏波後方散乱係数は、7.3.2 の 3) で示した通り、栽培形態別のモデルのうち直播栽培のモデルにおいては選択されたが、移植栽培のモデルには選択されなかった。以上のことから、モデルにおける VV 偏波後方散乱係数の選択は特に移植栽培の圃場においては有効でない可能性が示唆され、全データによるモデルでは VV 偏波後方散乱係数を除いて再度作成する必要があると考えられる。

#### 7.4 まとめ

本章では、C バンド 4 偏波 SAR である RADARSAT-2 データを使用し、稲作地帯における水稲の穂含水率の推定モデルの作成と検証を行った。衛星進行方向、入射角、栽培形態を示すダミー変数を加えることにより、モデルの精度を向上させることを目指した。まず、2016 年から 2019 年までの全 7 シーンの SAR データから 9 変数を算出し、重回帰分析により推定モデルの作成を行った。その結果、決定係数は低く RMSE も高かったことから、良好な精度を得られなかった。とくに、実測値が高い圃場では推定値が低く、過小評価されている傾向がみられ、実測値が低い圃場では推定値が高く、過大評価されている傾向がみられた。

次に、この結果を基準として、様々な条件で作成したモデルの精度について評価した。2016 年から 2019 年までの SAR データのうち、6 章と同様にアセンディング軌道で観測されたデータを使用したモデルと、ディセンディング軌道で観測されたデータを使用したモデルを作成した。また、入射角と栽培形態を示す各ダミー変数を追加した別の 2 モデルを作成し、精度の向上を検討した。その結果、アセンディング軌道のモデルで全データモデルよりも良好な決定係数が得られただけでなく、双方のモデルともダミー変数を加えることにより決定係数が更に向上した。とくにダミー変数を加えた場合は、アセンディング軌道のモデルで決定係数が 0.6 を上回り、高い精度が得られた。ダミー変数には双方のモデルとも栽培形態が選択されており、栽培形態の付加情報によって水稲の穂含水率がより高い精度で推定されることを示した。

この結果を踏まえ、2016 年から 2019 年までの SAR データを使用し、直播栽培の圃場のデータのみを使用したモデルと、移植栽培の圃場のみを使用したモデルを作成し、衛星進行

方向と入射角を示す各ダミー変数を加えて精度の向上を検討した。その結果、直播栽培のモデルに関しては0.3を超える決定係数が得られたが、ダミー変数を加えても決定係数の向上はみられなかった。ダミー変数には入射角と衛星進行方向の双方が選択されており、入射角や衛星進行方向の情報も推定モデルの精度を向上するために有効な変数であることが示唆された。一方、移植栽培のモデルに関しては、ダミー変数を加えることで精度の向上はみられたものの、決定係数は0.3を下回り、精度の高いモデルの作成は困難であった。こちらのモデルにもダミー変数として入射角が選択された。

最後のモデルの作成として、2016年から2019年までの全7シーンのSARデータから算出された9変数と、全3種類のダミー変数のすべてを使用してモデルを作成した。その結果、決定係数は0.452、RMSEは6.48%と、ダミー変数を使用していないモデルと比較して精度が向上した。ダミー変数を使用することにより、これまでみられていた過大評価や過小評価の傾向が緩和された。この結果は、ダミー変数を含めたアセンディング軌道によるモデルと比較して精度は低いものの、ダミー変数を含めたディセンディング軌道によるモデルおよび直播栽培によるモデル、移植栽培によるモデルと比較すると高い精度が得られた。したがって、栽培形態ごとにモデルを作成しなくても1つのモデルにまとめることができることが示された。

このモデルには、HV偏波後方散乱係数、VV偏波後方散乱係数、エントロピ、栽培形態のダミー変数の4種類の変数が選択された。これまでのほぼすべてのモデルにおいてエントロピが選択されていた。エントロピは散乱の形態の複雑さを示す変数であり、穂の水分の差異によって散乱形態が変化する様相を補足することができたと考えられる。また、作物体の物理量の差異で値が変化しやすい体積散乱またはHV偏波後方散乱係数もいくつかのモデルで選択される傾向がみられ、これらが選択された場合に標準偏回帰係数は高い値となった。本モデルでもHV偏波後方散乱係数は同様に高い標準偏回帰係数が得られた。したがって、HV偏波後方散乱係数やエントロピは特に含水率を推定するための重要な要素となったと考えられる。ダミー変数では栽培形態が選択されており、衛星データの入射角や衛星進行方向よりも水稻の種類がモデルの精度を左右する要因となることが示された。これは、水稻の栽培形態ごとの成長量や成長時期、生育密度の違いがモデルに影響する可能性が示唆される。

以上の結果より、水稻の穂含水率を推定するモデルを作成するにあたり、ダミー変数の有

効性が示された点や、ダミー変数を使用することでモデルを分割する必要性が低くなった点から、全データにダミー変数を加えたモデルが最も有効な水稻の穂含水率の推定モデルであると考えられる。

この全データにダミー変数を加えたモデルについて、交差検証を行った。交差検証には、5分割による検証手法を用いた。モデル作成に当たっては、当該のモデルにおいて選択された4変数のみを使用したダミー変数あり4変数モデルと、すべての変数で再度ステップワイズをし直したダミー変数あり再構築モデル、ダミー変数を使用していないダミー変数なしモデルの3種類について、それぞれ交差検証を行った。その結果、作成されたモデルの決定係数、RMSEの平均はダミー変数あり再構築モデルの方が高い精度であったものの、検証結果におけるRMSEの平均はダミー変数あり4変数モデルの方が低く、良好な精度となった。以上の結果より、HV偏波後方散乱係数、VV偏波後方散乱係数、エントロピ、栽培形態の4変数を使用したモデルは汎化性能が高く、最も有効なモデルであると結論づけた。

最後に、全データにダミー変数を加えたモデルをもとに、地域全体の水稻の穂含水率の予測値を示す推定マップを作成した。2019年の2農業法人におけるすべての水稻作付圃場の区画を示すポリゴンを用い、2019年の2シーンのSARデータより各圃場の平均値を算出し、このモデルに適用した結果を推定マップに出力した。その結果、2シーンそれぞれで含水率が低い圃場は一致しているものの、経時変化で穂含水率が上昇している傾向がみられた。とくに8月20日のマップでは、6章で示した現地調査の時期の含水率より低い値となった。この要因として、VV偏波後方散乱係数が低い値であったことが要因として考えられる。以上のことから、VV偏波後方散乱係数を除いて再度モデルを作成する必要性が示唆された。

## 8 結言

本研究では、合成開口レーダ (SAR) によるリモートセンシングを中心に利用して、畑作地帯における作付作物の分類や、稲作地帯における水稲生育状況の推定への適用可能性を評価することを目的とした。

畑作地帯における作付作物の分類として、1) X バンド SAR データを利用した分類、2) C バンド SAR データを利用した分類および両者を併用した分類、3) 機械学習による分類アルゴリズムの精度比較や SAR データ数を削減した効率的な分類、4) C バンド SAR データおよび L バンド SAR データを併用した効率的な分類の 4 項目を遂行した。

さらに、稲作地帯における水稲生育状況の推定として、4) C バンド SAR データおよび高解像度光学センサを利用した水稲の穂含水率の推定モデル作成、5) C バンド SAR データおよびダミー変数を利用した水稲の穂含水率の推定モデル作成の 2 項目を遂行した。

以上の 6 項目について、以下に結果を列挙した。

1) X バンドの SAR を搭載した TerraSAR-X の 2 偏波データから、後方散乱係数 (HH 偏波, VV 偏波) と固有値解析法による散乱成分 (エントロピ, アルファ角) の計 4 変数を算出した。これらの 4 変数から、機械学習アルゴリズムの Random Forests を使用し、北海道の十勝地方における代表的な畑作 6 作物 (小麦, 豆類, 馬鈴薯, 甜菜, 牧草, トウモロコシ) を対象に、圃場単位での作付作物の分類を行った。その結果、6 月から 7 月にかけて観測された 5 時期のデータを用いることにより、全体精度が 0.895 となった。また、データ取得コストの削減を目的として、時期数を削減して分類を行った。その結果、2 時期のデータを使用することで 0.8 を超える全体精度が得られ、4 時期のデータを使用することで 5 時期での分類と有意な差のない結果が得られた。

2) X バンドの SAR データより高い精度での作付作物の分類を目的として、C バンドの SAR を搭載した RADARSAT-2 の 4 偏波データを使用し、Random Forests から 6 作物における圃場単位での作付作物の分類を実行した。4 偏波 SAR データから 16 変数を算出し、最も高い精度が得られる変数の組み合わせを検討した。その結果、後方散乱係数 (HH 偏波, HV 偏波, VV 偏波) と Van Zyl の散乱モデル分解法による散乱成分 (表面散乱, 体積散乱, 2 回散乱), 固有値解析法による散乱成分 (エントロピ, アニソトロピ, アルファ角) の計 9 変数を用いることにより、5 時期のデータから 0.922 の全体精度が得られた。また、RADARSAT-

2 の 9 変数と TerraSAR-X の 4 変数を併用することにより、各 5 時期のデータから 0.934 の全体精度が得られた。最後に、データ取得コストの削減を目的として、時期数を削減して分類を行った。その結果、RADARSAT-2 のみでの分類では、3 時期のデータを使用することで 0.9 を超える全体精度が得られ、4 時期のデータを使用することで 5 時期での分類と有意な差のない結果が得られた。RADARSAT-2 と TerraSAR-X を併用した分類では、2 時期のデータを使用することで 0.9 を超える全体精度が得られ、3 時期のデータを使用することで 5 時期での分類と有意な差のない結果が得られた。

3) Random Forests 以外の決定木による機械学習アルゴリズムのうち、SAR による分類で使用された例がない機械学習アルゴリズム (Extremely Randomized Trees および Random Ferns) を使用し、高精度化や効率化に関する検討を行った。RADARSAT-2 データからは 9 変数を算出した。TerraSAR-X データからは、後方散乱係数 (HH 偏波, VV 偏波) と固有値解析法による散乱成分 (エントロピ, アルファ角) のほか、Raney の散乱モデル分解法による散乱成分 (表面散乱, 体積散乱, 2 回散乱) を加えた計 7 変数を算出した。これらの変数を使用し、6 作物における圃場単位での作付作物の分類に、3 種類の機械学習アルゴリズムを適用した。その結果、Extremely Randomized Trees はほぼ全ての教師データ比率や時期数の場合において、Random Forests より高い全体精度が得られた。この Extremely Randomized Trees を使用して、TerraSAR-X データと RADARSAT-2 データの分類における精度や特徴を比較した。その結果、1 時期の SAR での分類では、7 月までは RADARSAT-2 データの方が概ね高い全体精度を得られ、8 月は TerraSAR-X データの方が高い全体精度を得られる傾向がみられた。多時期での分類では、ほぼすべての時期数で RADARSAT-2 データの方が高い全体精度を得られた。最後に、少ないデータ数による効率的な分類を検討したところ、7 月上旬までの RADARSAT-2 データ 2 時期と 7 月中旬以降の TerraSAR-X データ 4 時期を併用することで、0.953 の全体精度が得られた。この結果は、両者を各 6 時期全て使用した分類と有意差がなかった。

4) RADARSAT-2 データや TerraSAR-X データと比較して安価で購入できる、L バンドの SAR を搭載した ALOS-2/PALSAR-2 データを使用し、コストを削減した作付作物の分類手法について提示した。RADARSAT-2 の 4 偏波データと、PALSAR-2 の単偏波データおよび 4 偏波データを使用し、Random Forests を用いて 6 作物における圃場単位での分類を実行した。単偏波データからは後方散乱係数 (HH 偏波) を、4 偏波データからは 9 変数を算出し

た。1 時期の 4 偏波データで RADARSAT-2 データと PALSAR-2 データの分類結果を比較したところ、PALSAR-2 データからは、RADARSAT-2 データより良好な全体精度を得られなかった。3 時期の RADARSAT-2 データと PALSAR-2 データで分類結果を比較したところ、RADARSAT-2 データによる分類の方が全体精度は高かったが、両者を併用することにより全体精度が向上した。また、4 時期の RADARSAT-2 データと 2 時期の PALSAR-2 データを併用した分類では、0.9 を超える全体精度が得られた。

5) C バンドの SAR を搭載した RADARSAT-2 の 4 偏波データから 9 変数を算出し、高解像度光学センサである Geosyde-1 データから NDVI を算出した。これらの変数を用いて、ステップワイズの変数減少法を用いた重回帰分析を使用し、北海道の空知地方における代表先な作物である水稲を対象に、圃場単位での穂含水率の推定を行った。まず、RADARSAT-2 データによるモデルの作成を行った。モデルの作成にあたっては、衛星の進行方向や栽培形態が精度に影響を及ぼすと考え、全データを使用したモデルの他に、衛星の進行方向ごとのモデルや栽培形態ごとのモデルを作成した。その結果、全データを使用したモデルの決定係数は 0.152 であったが、衛星進行方向ごと、栽培形態ごとのモデルはそれぞれ決定係数の向上がみられた。とくにアセンディング軌道によるモデルと移植栽培によるモデルの決定係数は 0.5 を超えた。次に、Geosyde-1 データを使用してモデルを作成した。Geosyde-1 データのみを使用したモデルでは、湛水していない圃場のみを対象とすることで、0.923 の決定係数が得られた。RADARSAT-2 データと Geosyde-1 データを併用したモデルでは、さらに精度が向上し、0.965 の決定係数が得られた。最後に、RADARSAT-2 データより作成した 5 種類のモデルについて、翌年度のデータを用いて検証を行った。全データを使用したモデルは RMSE が最も低く、6%程度となった。

6) C バンドの SAR を搭載した RADARSAT-2 の 4 偏波データから算出された 9 変数と、3 種類のダミー変数 (衛星の進行方向、衛星の入射角、水稲の栽培形態) を使用し、圃場単位での水稲の穂含水率を推定するモデルを作成した。モデル作成には、ステップワイズの変数減少法を用いた重回帰分析を使用した。全データを使用したモデル、衛星進行方向ごとのモデル、栽培形態ごとのモデルのいずれにおいても、ダミー変数を使用することで精度の向上がみられた。とくに、全データにダミー変数を加えたモデルは、0.452 の決定係数が得られた。この結果は、ダミー変数を含めたディセンディング軌道によるモデルおよび直播栽培のモデル、移植栽培のモデルより高い精度であった。最後に、この全データにダミー変数を加

えたモデルについて、5分割による交差検証を行った。その結果、検証結果における RMSE は 6.70%と、モデル作成における RMSE と同等の精度となった。モデルには、HV 偏波、VV 偏波、エントロピ、栽培形態の 4 変数を選択することで、特に低い RMSE が得られた。

以上のことから、合成開口レーダ (SAR) により取得された衛星データは、畑作地帯における作付作物の分類や稲作地帯における水稻の穂含水率の推定について、高い精度での適用が可能であった。作付作物の分類では、7月上旬までの C バンド SAR と 7月中旬以降の X バンド SAR を併用し、各データから後方散乱係数、散乱モデル分解法による散乱成分、固有値解析法による散乱成分を算出して、ERT のアルゴリズムへ適用することが特に有効であることが明らかとなった。水稻の穂含水率の推定では、8月下旬の C バンド SAR を使用し、体積散乱とエントロピを算出し、栽培体系のダミー変数を用いることで、有効な重回帰モデルを作成できることを示した。これらの結果から、合成開口レーダは農業分野において圃場の情報を取得するための有望な手段となることが結論づけられた。

## 謝辞

本研究を遂行するにあたり、北海道大学大学院農学研究院の谷宏准教授には、テーマの選定から研究計画、現地調査、論文上梓に至るまで、主査として多大なご指導、ご鞭撻を賜った。また、ご多忙にもかかわらず、井上京教授、鮫島良次教授には副査として本論文の校閲を頂き、貴重な助言を賜った。また、王秀峰准教授、山本忠男講師には、本研究を遂行するにあたり平素より貴重な助言を賜った。ここに厚く感謝申し上げる。

研究に使用した衛星データとして、RADARSAT-2 データの一部および TerraSAR-X データについては、株式会社パスコより提供して頂いたうえで、望月貫一郎氏、木村篤史氏には衛星データの処理や研究への利用法に関する有益な情報を賜った。ALOS-2/PALSAR-2 データについては、宇宙航空研究開発機構（JAXA）の公募型共同研究（RA-6 PI-3288）により提供を受けたものである。作付作物分類に用いた対象圃場に関する GIS データは、十勝農業共済組合から提供して頂いた。ここに深く御礼申し上げます。

静岡大学農学部の菌部礼助教には、ソフトウェアの使用法や研究の方向性に関して多大な助言を賜った。株式会社スマートリンク北海道の小林伸行氏、西飯弘行氏、古村祐基氏には、現地調査に同行して頂き、調査の方法および圃場の計測等に関して多大な助言を賜った。佐藤農場の佐藤裕一氏、堀川農場の堀川寛氏、坂東農場の坂東俊徳氏をはじめ、清水町、芽室町、幕別町、岩見沢市の各農業法人の皆様には、計測を行うために圃場を調査地として提供して頂いた上で、これらの調査圃場に関する有益な情報を賜った。

研究を遂行するにあたり、農業土木学研究室および生物環境情報学研究室の学生、卒業生および関係者各位には衛星データの解析や現地調査、各種データの作成等で多大な助力を頂いた。記して謝意を表する次第である。

最後に、本研究は JSPS 科研費 JP18J20030 の助成を受けたものである。



## 引用文献

- Airbus Defence and Space (2015) Radiometric Calibration of TerraSAR-X Data.  
<[http://www.geo-airbusds.com/files/pmedia/public/r465\\_9\\_tsx-x-itd-tn-0049-radiometric\\_calculations\\_i3.00.pdf](http://www.geo-airbusds.com/files/pmedia/public/r465_9_tsx-x-itd-tn-0049-radiometric_calculations_i3.00.pdf)>, 01/06/2016 referred.
- Bindlish, R. and A. P. Barros. (2001) Parameterization of vegetation backscatter in radar-based, soil moisture estimation. *Remote Sensing of Environment*, **76**, pp.130-137.
- Blaes, X., L. Vanhalle and P. Defourny. (2005) Efficiency of crop identification based on optical and SAR image time series. *Remote Sensing of Environment*, **96** (3-4), pp. 352-365.
- Breiman, L. (2001) Random forests. *Machine Learning*, **45** (1), pp. 5-32.
- Cloude, S. R. (2007) The dual-polarization entropy/alpha decomposition: a pulsar case study, *Proc. POLInSAR*, Frascati, Italy, pp. 1-6.
- Cloude, S. R. and Pottier, E. (1996) A review of target decomposition theorems in radar polarimetry. *IEEE Transactions on Geoscience and Remote Sensing*, **34** (2), pp. 498-518.
- Cohen, J. (1960) A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, **20**, pp. 37-46.
- Congalton, R. G. and K. Green. (2008) *Assessing the accuracy of remotely sensed data: Principles and practices*. Boca Raton, Florida, United States: CRC Press, pp. 1-183.
- Englhart, S., V. Keuck and F. Siegert. (2011) Aboveground biomass retrieval in tropical forests - the potential of combined X- and L-band SAR data use. *Remote Sensing of Environment*, **115** (5), pp. 1260-1271.
- ESA (2015) ESA science toolbox exploitation platform. <<http://step.esa.int/main/>>, 01/06/2016 referred.
- Freeman, A. and Durden, S. L. (1998) A three-component scattering model for polarimetric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, **36** (3), pp. 963-973.
- 古川猛, 古関潤一 (2019) 干渉 SAR を利用した地盤変状のモニタリング. 土と岩, **67** (2), pp. 4-11.
- Gens, R., D. Atwood and E. Pottier (2013) Geocoding of polarimetric processing results: Alternative processing strategies. *Remote Sensing Letters*, **4** (1), pp. 39-45.

- Geurts, P., Ernst, D. and Wehenke, L. (2006) Extremely randomized trees. *Machine learning*, **63** (1), pp. 3-42.
- Hartfield, K. A., Marsh, S. E., Kirk, C. D. and Carriere, Y. (2013) Contemporary and historical classification of crop types in Arizona. *International Journal of Remote Sensing*, **34** (17), pp. 6024-6036.
- Huang, X., Wang, J., Shang, J., Liao, C. and Liu, J. (2017) Application of polarization signature to land cover scattering mechanism analysis and classification using multi-temporal C-band polarimetric RADARSAT-2 imagery. *Remote Sensing of Environment*, **193**, pp. 11-28.
- Immitzer, M., Vuolo, F. and Atzberger, C. (2016) First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe. *Remote Sensing*, **8** (3), 166.
- 石塚直樹 (2016) マイクロ波合成開口レーダを用いた農地計測事例. 計測と制御, **55** (9), pp. 814-817.
- 石塚直樹, 大内和夫 (2017) 合成開口レーダの農業への応用. 日本リモートセンシング学会誌, **37** (3), pp. 182-192.
- 井上吉雄 (2016) 食糧安全保障と農業生産の情報化に向けた先進リモートセンシング技術. 計測と制御, **55** (9), pp. 747-749.
- Jiao, X. F., Kovacs, J. M., Shang, J. L., McNairn, H., Walters, D., Ma, B. L. and Geng, X. Y. (2014) Object-oriented crop mapping and monitoring using multi-temporal polarimetric RADARSAT-2 data. *ISPRS Journal of Photogrammetry and Remote Sensing*, **96**, pp. 38-46.
- Ju, W., P. Gao, J. Wang, Y. Zhou and X. Zhang (2010) Combining an ecological model with remote sensing and GIS techniques to monitor soil water content of croplands with a monsoon climate. *Agricultural Water Management*, **97** (8), pp. 1221-1231.
- 国土地理院 (2015) 基盤地図情報ダウンロードサービス.  
<<http://fgd.gsi.go.jp/download>>, 05/11/2015 referred.
- Kim, N., Ha, K. J., Park, N. W., Cho, J., Hong, S. and Lee, Y. W. (2019) A Comparison Between Major Artificial Intelligence Models for Crop Yield Prediction: Case Study of the Midwestern United States, 2006-2015. *ISPRS International Journal of Geo-Information*, **8** (5), 240.
- 高度情報通信ネットワーク社会推進戦略本部 (2014) 農業情報創成・流通促進戦略.  
<[http://www.kantei.go.jp/jp/singi/it2/kettei/pdf/senryakuzenbun\\_140603.pdf](http://www.kantei.go.jp/jp/singi/it2/kettei/pdf/senryakuzenbun_140603.pdf)>, 22/03/2020

referred.

- Krahwinkler, P. and J. Rossmann. (2013) Tree species classification and input data evaluation. *European Journal of Remote Sensing*, **46**, pp. 535-549.
- Kucuk, C., Taskin, G. and Erten, E. (2016) Paddy-Rice Phenology Classification Based on Machine-Learning Methods Using Multitemporal Co-Polar X-Band SAR Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **9** (6), pp. 2509–2519.
- Kursa M. B. (2014) rFerns: An Implementation of the Random Ferns Method for General-Purpose Machine Learning. *Journal of Statistical Software*, **61** (10), pp. 1-13.
- Lee, J. S., Grunes, M. R. and de Grandi, G. (1999) Polarimetric SAR Speckle filtering and its implication for classification. *IEEE Transactions on Geoscience and Remote Sensing*, **37**, pp. 2363–2373.
- Lee, J. S., M. R. Grunes and E. Pottier. (2001) Quantitative comparison of classification capability: Fully polarimetric versus dual and single-polarization SAR. *IEEE Transactions on Geoscience and Remote Sensing*, **39** (11), pp. 2343–2351.
- Liaw, A. and M. Wiener. (2002) Classification and Regression by Random Forest. *R News*, **2** (3), pp. 18–22.
- Mascolo, L., Forino, G., Nunziata, F., Pugliano, G. and Migliaccio, M. (2019) A New Methodology for Rice Area Monitoring With COSMO-SkyMed HH–VV PingPong Mode SAR Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **12** (4), pp. 1076-1084.
- McNairn, H., Ellis, J., Van Der Sanden, J. J., Hirose, T. and Brown, R. J. (2002) Providing crop information using RADARSAT-1 and satellite optical imagery. *International Journal of Remote Sensing*, **23** (5), pp. 851-870.
- McNairn, H., J. L. Shang, X. F. Jiao and C. Champagne. (2009) The contribution of ALOS PALSAR multipolarization and polarimetric data to crop classification. *IEEE Transactions on Geoscience and Remote Sensing*, **47** (12), pp. 3981–3992.
- McNemar, Q. (1947) Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, **12**, pp. 153-157.
- MDA (2004) RADARSAT-1 Data Products Specifications.

- <[http://gs.mdacorporation.com/includes/documents/R1\\_PROD\\_SPEC.pdf](http://gs.mdacorporation.com/includes/documents/R1_PROD_SPEC.pdf)>, 01/06/2016 referred.
- 水野正樹, 神山嬢子, 江川真史, 佐藤匠, 蒲原潤一 (2014) 2 偏波 SAR 画像による大規模崩壊および河道閉塞箇所の判読調査手法 (案). 国土技術政策総合研究所資料, 791.
- Ndikumana, E., D. Ho Tong Minh, H. Thu Dang Nguyen, N. Baghdadi, D. Courault, L. Hossard and I. El Moussawi (2018) Estimation of Rice Height and Biomass Using Multitemporal SAR Sentinel-1 for Camargue, Southern France. *Remote Sensing*, **10** (8), 1394.
- NGA GEOnet Names Server (2013) Earth Gravitational Model 2008 (EGM2008) <<http://earth-info.nga.mil/GandG/wgs84/gravitymod/egm2008/>>, 05/12/2015 referred.
- Nichol, J. E. and Sarker, M. L. R. (2011) Improved Biomass Estimation Using the Texture Parameters of Two High-Resolution Optical Sensors. *IEEE Transactions on Geoscience and Remote Sensing*, **49** (3), pp. 930-948.
- 日本リモートセンシング学会 (2011) 基礎からわかるリモートセンシング. 理工図書.
- 日本リモートセンシング研究会 (2001) 図解リモートセンシング. 社団法人日本測量学会.
- 野口伸 (2016) ICT 農業とリモートセンシング. 日本ロボット学会誌, **34** (2), pp. 100-102.
- Nordkvist, K., A. Granholm, J. Holmgren, H. Olssona and M. Nilsson. (2012) Combining optical satellite data and airborne laser scanner data for vegetation classification. *Remote Sensing Letters*, **3** (5), pp. 393-401.
- 農林水産省 (2020) 面積調査の概要.  
<<https://www.maff.go.jp/j/tokei/kouhyou/sakumotu/menseki/gaiyou/index.html#1>>,  
10/12/2020 referred.
- Oguro, Y., Suga, Y., Takeuchi, S., Ogawa, H. and Tsuchiya, K. (2003) Monitoring of a rice field using landsat-5 TM and landsat-7 ETM+ data. *Advances in Space Research*, **32** (11), pp. 2223-2228.
- Ozuysal, M., Calonder, M., Lepetit, V. and Fua, P. (2010) Fast keypoint recognition using random ferns. *IEEE transactions on pattern analysis and machine intelligence*, **32** (3), pp. 448-461.
- Parisot P. and Vleeschouwer, C. D. (2017) Scene-specific classifier for effective and efficient team sport players detection from a single calibrated camera. *Computer Vision and Image Understanding*, **159**, pp. 74-88.

- Pottier, E. and Ferro-Famil, L. (2012) PolSARPro V5.0: An ESA educational toolbox used for self-education in the field of POLSAR and POL-INSAR data analysis. *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*, pp. 7377-7380.
- Qi, Z. X., Yeh, A. G. O. and Li, X. (2019) Scattering-Mechanism-Based Investigation of Optimal Combinations of Polarimetric SAR Frequency Bands for Land Cover Classification. *Photogrammetric Engineering and Remote Sensing*, **85** (11), pp. 799-813.
- R Core Team (2016 updated) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <<http://www.R-project.org/>>, 01/06/2016 referred.
- Raney, R.K. (2013) M-Chi Decomposition of Imperfect Hybrid Dual-Polarimetric Radar Data. *Proceedings of 6th International Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry: PolInSAR 2013*.
- Ramadhani, F., Pullanagari, R., Kereszturi, G. and Procter, J. (2020) Automatic Mapping of Rice Growth Stages Using the Integration of SENTINEL-2, MOD13Q1, and SENTINEL-1. *Remote Sensing*, **12** (21), 3613.
- 齋藤元也, 石塚直樹, 坂本利弘 (2017) 日本における農業リモートセンシング研究の軌跡. *日本リモートセンシング学会誌*, **37** (3), pp. 193-203.
- 境谷栄二, 井上吉雄 (2013) 米の適期収穫への航空機および衛星リモートセンシングの実践的利用. *日本リモートセンシング学会誌*, **33** (3), pp. 185-189.
- Sarker, L. and Nichol, J. (2011) Improved forest biomass estimates using ALOS AVNIR-2 texture indices. *Remote Sensing of Environment*, **115** (4), pp. 968-977.
- 佐々木次郎, 伊藤修 (2000) 水稻籾水分を利用した刈取適期の予測.  
<<https://www.naro.affrc.go.jp/org/tarc/seika/jyouhou/H12/tnaes00013.html>>, 20/11/2020 referred.
- Sharifi, A. and Hosseingholizadeh, M. (2019) Application of Sentinel-1 Data to Estimate Height and Biomass of Rice Crop in Astaneh-ye Ashrafiyeh, Iran. *Journal of the Indian Society of Remote Sensing*, **48** (4), pp. 11-19.
- 志賀弘行 (2010) 農業分野における衛星画像の利用.  
<[http://www.alic.go.jp/joho-s/joho07\\_000102.html](http://www.alic.go.jp/joho-s/joho07_000102.html)>, 20/12/2017 referred.

- Sicre, C. M., Fieuzal, R. and Baup, F. (2020) Contribution of multispectral (optical and radar) satellite images to the classification of agricultural surfaces. *International Journal of Applied Earth Observation and Geoinformation*, **84**, 101972.
- Simm J., de Abril I. and Sugiyama M. (2014) Tree-Based Ensemble Multi-Task Learning Method for Classification and Regression. *IEICE Transactions on Information and Systems*, **97** (6), pp. 1677-1681.
- Skakun, S., Kussul, N., Shelestov, A. Y., Lavreniuk, M. and Kussul, O. (2015) Efficiency Assessment of Multitemporal C-Band Radarsat-2 Intensity and Landsat-8 Surface Reflectance Satellite Imagery for Crop Classification in Ukraine. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **9** (8), pp. 3712-3719.
- Soltaninejad, M., Yang, G., Lambrou, T., Allinson, N., Jones, T. L., Barrick, T. R., Howe, F. A. and Ye, X. J. (2017) Automated brain tumour detection and segmentation using superpixel-based extremely randomized trees in FLAIR MRI. *International journal of computer assisted radiology and surgery*, **12** (2), pp. 183-203.
- Son, N. T., Chen, C. F., Chen, C. R. and Minh, V. Q. (2018) Assessment of Sentinel-1A data for rice crop classification using random forests and support vector machines. *Geocarto International*, **33** (6), pp. 1-32.
- Sonobe R., Tani H., Wang X., Kobayashi N. and Shimamura H. (2014a) Random forest classification of crop type using multi-temporal TerraSAR-X dual-polarimetric data. *Remote Sensing Letters*, **5** (2), pp. 157-164.
- Sonobe R., Tani H., Wang X., Kobayashi N. and Shimamura H. (2014b) Parameter tuning in the support vector machine and random forest and their performances in cross- and same-year crop classification using TerraSAR-X. *International Journal of Remote Sensing*, **35** (23), pp. 7898-7909.
- Sonobe R., Tani H., Wang X., Kobayashi N. and Shimamura H. (2014c) Winter Wheat Growth Monitoring Using Multi-temporal TerraSAR-X Dual-polarimetric Data. *Japan Agricultural Research Quarterly*, **48** (4), pp. 471-476.
- Sonobe, R., Yamaya, Y., Tani, H., Wang, X., Kobayashi, N. and Mochizuki, K. (2018) Crop classification from Sentinel-2 derived vegetation indices using ensemble learning. *The*

- Journal of Applied Remote Sensing*, **12** (2), 026019.
- Sonobe, R. (2019) Parcel-Based Crop Classification Using Multi-Temporal TerraSAR-X Dual Polarimetric Data. *Remote Sensing*, **11** (10), 1148.
- Srikanth, P., Ramana, K. V., Deepika, U., Chakravarthi, P. K. and Sai, M. V. R. S. (2016) Comparison of Various Polarimetric Decomposition Techniques for Crop Classification. *Journal of the Indian Society of Remote Sensing*, **44** (4), pp. 635–642.
- Torbick, N., Salas, W., Xiao, X. M., Ingraham, P., Fearon, M. G., Biradar, C., Zhao, D. L., Liu, Y., Li, P., Zhao, Y. L. (2011) Integrating SAR and optical imagery for regional mapping of paddy rice attributes in the Poyang Lake Watershed, China. *Canadian Journal of Remote Sensing*, **37** (1), pp. 17-26.
- 辻野照久 (2014) 衛星画像を利用した農業生産統計. 科学技術動向, **145**, pp. 26-30.
- Valcarce, R., Juan D., Sanchez, M. L., Sanchez, N., Perez, B. A. and Fernandez, J. M. (2019) Influence of Incidence Angle in the Correlation of C-band Polarimetric Parameters with Biophysical Variables of Rainfed Crops. *Canadian Journal of Remote Sensing*, **44** (6), pp. 643-659.
- Van Zyl, J. J. (1990) Calibration of polarimetric radar images using only image parameters and trihedral corner reflector responses. *IEEE Transactions on Geoscience and Remote Sensing*, **28** (3), pp. 337-348.
- Van Zyl, J. J., Arii, M. and Kim, Y. (2011) Model-Based Decomposition of Polarimetric SAR Covariance Matrices Constrained for Nonnegative Eigenvalues. *IEEE Transactions on Geoscience and Remote Sensing*, **49** (9), pp. 3452-3459.
- Whelen, T. and Siqueira, P. (2018) Coefficient of variation for use in crop area classification across multiple climates. *International Journal of Applied Earth Observation and Geoinformation*, **67**, pp.114-122.
- Xie, Q., Wang, J., Liao, C., Shang, J., Juan, M. L. S., Fu, H. and Liu, X. (2019) On the Use of Neumann Decomposition for Crop Classification Using Multi-Temporal RADARSAT-2 Polarimetric SAR Data. *Remote Sensing*, **11** (7), 776.
- Xu, J., Li, Z., Tian, B. S., Huang, L., Chen, Q. and Fu, S. T. (2014) Polarimetric analysis of multi-temporal RADARSAT-2 SAR images for wheat monitoring and mapping. *International Journal of Remote Sensing*, **35** (10), pp. 3840-3858.

- Yaghouti, H., Pazira, E., Amiri, E. and Masihabadi, M. (2019) The Feasibility of Using Vegetation Indices and Soil Texture to Predict Rice Yield. *Polish Journal of Environmental Studies*, **28** (4), pp. 2473-2481.
- 山口芳雄 (2007) レーダーポ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 –. 社団法人電子情報通信学会.
- Yamaguchi, Y., Moriyama, T., Ishido, M. and Yamada, H. (2005) Four-component scattering model for polarimetric SAR image decomposition. *IEEE Transactions on Geoscience and Remote Sensing*, **43** (8), pp. 1699-1706.
- Zhang, D., Fang, S., She, B., Zhang, H., Jin, N., Xia, H., Yang, Y. and Ding, Y. (2019) Winter Wheat Mapping Based on Sentinel-2 Data in Heterogeneous Planting Conditions. *Remote Sensing*, **11**(22), 2647.
- Zhao, L. L., Yang, J., Li, P. X. and Zhang, L. P. (2014) Characteristics Analysis and Classification of Crop Harvest Patterns by Exploiting High-Frequency Multi Polarization SAR Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **7** (9), pp. 3773-3783.
- Zhong, L. H., Gong, P. and Biging, G. S. (2014) Efficient corn and soybean mapping with temporal extendability: A multi-year experiment using Landsat imagery. *Remote Sensing of Environment*, **140**, pp. 1-13.