

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圃場の作物分類

Crop Classification by Machine Learning Algorithm with Combination of X- and C-band SAR Data

山谷 祐貴*・谷 宏**・王 秀峰**・菌部 礼***
小林 伸行****・望月貫一郎*****・野田 萌*****

*Yuki YAMAYA, Hiroshi TANI, Xiufeng WANG, Rei SONOBE
Nobuyuki KOBAYASHI, Kan-ichiro MOCHIZUKI and Megumi NODA*

Abstract : A crop classification method using satellite data is proposed as an alternative to the existing ground survey. In this study, crop types were classified using two kinds of SAR data (i.e., TerraSAR-X X-band dual-polarization data and Radarsat-2 C-band fully-polarization data) and Random Forests. Sigma naught polarimetric parameters were calculated from SAR data and classifications were conducted using the following four different datasets; Case 1: all parameters calculated from Radarsat-2, Case 2: all parameters calculated from Radarsat-2 and sigma naught calculated from TerraSAR-X data, Case 3: all parameters calculated from Radarsat-2 and polarimetric parameters calculated from TerraSAR-X data, and Case 4: all parameters calculated from Radarsat-2 and both sigma naught and polarimetric parameters calculated from TerraSAR-X. The highest overall accuracy of 0.934 was achieved by Case 4, and there were significant differences with the other classification results ($p > 0.05$, based on Z-test). These results reveal that combining two kinds of SAR data can be improved classification accuracy.

1. 背景と目的

農業地域における作付状況に関する情報は、作物ごとの農地面積の調査や収量の予測、災害発生時の被害把握など、様々な用途に供せられている。よって、作物の作付状況をより広範囲で把握することは重要である。しかし、作付作物の分類の詳細を継続的かつ長期的、広域的に把握していくためには、現在人の手によって行われているような現地調査では困難となる。したがって、現地調査に代わる手法の確立は必要不可欠であり、その中でも衛星リモートセンシング技術は、農

業分野において特に有望な情報収集手段の一つになりうる。

本研究では、衛星リモートセンシングの中でも、マイクロ波を使用した能動型センサである合成開口レーダ (SAR, Synthetic Aperture Radar) を用いた技術を利用した。SAR は昼夜問わず安定した観測が可能であるだけでなく、大気や天候の影響を受けずにデータを取得することが可能である (Bindlish and Barros, 2001)。したがって、SAR データを用いた作物のモニタリングは、非常に汎用性、堅牢性が高く、毎年安定して確実に結果を得られる手法であることから (石塚, 2016)、多時期のデータを使用した作付作物の分類にも効果的であると考えられる。

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.*, 2014)、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ている (山谷ほか, 2017a)。また、Cバンド SAR データを使用した作付け作物の分類でも、同様の方法により 0.922 というさ

* 北海道大学大学院農学院

** 北海道大学大学院農学研究院

*** 静岡大学農学部

**** 株式会社スマートリンク北海道

***** 株式会社バスコ

***** 札幌市役所

「写真測量とリモートセンシング」VOL. 57, NO. 2, 2018

らに高い全体精度が得られている(山谷ほか, 2017b)。そこで, 本研究では5時期で高い全体精度が得られたCバンドSARデータをベースにXバンドSARデータを併用し, 畑作地帯における作付作物の分類を行った。その結果から, 1種類のバンドのみで分類を行う場合より高い精度で分類が行えるかどうかを評価した。分類の手法には, 上記の先行研究と同様に, パターン認識において特に高い精度が期待できる機械学習アルゴリズムの1つであるランダムフォレストを使用した。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だけでなく, 各種の散乱成分分解から算出される散乱成分も変数として使用した。

2. 研究方法

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1)。

対象作物は, 当該地域における主要作物である小麦, 豆類, 馬鈴薯, 甜菜, 牧草, トウモロコシの6種類とし, 全5,076圃場を解析対象とした。なお, 豆類には小豆・大豆・菜豆を含み, 牧草にはチモシー・オーチャードグラスを, トウモロコシには食用・飼料用をそれぞれ含んでいる。

2.2 使用した衛星データ

本研究では, CバンドSARを搭載したRADARSAT-2の4偏波データおよびXバンドSARを搭載したTerraSAR-Xの2重偏波データをそれぞれ5時

表1 各衛星データの取得時期

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

期使用した。表1に, それぞれの衛星データの取得時期を示す。両者の衛星データは撮像された時期が近くなるようにそれぞれ選定した。データはシングルルックコンプレックス(SLC)であるため, 作付作物分類を実施するにあたり, 前処理として, スペックルフィルタリング処理やマルチルック処理, オルソ補正を行った。なお, オルソ補正には, 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。

その後RADARSAT-2データからは, 各偏波の後方散乱係数を式(1)から算出した。

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

ここで, θ_{loc} は衛星の入射角を示し, 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ル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, モノスタティックでの計測の場合は値が理論上等しくなるため(Jiao *et al.*, 2014), RADARSAT-2データより算出される σ_{VV}^0 は使用せず, 残りの3偏波の後方散乱係数(σ_{HH}^0 , σ_{HV}^0 , σ_{VH}^0)を使用した(表2)。また3偏波の後

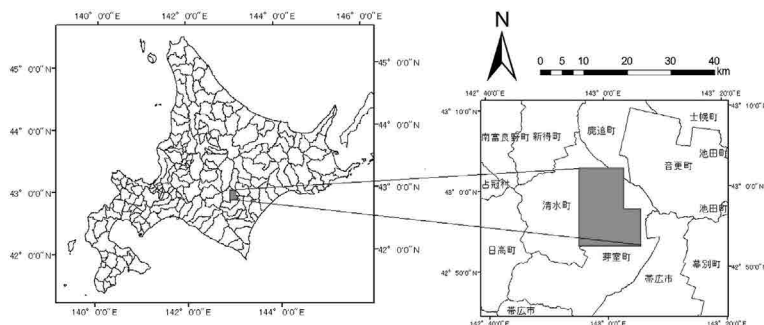


図1 対象地域

表2 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		
RADARSAT-2	後方散乱係数	HH 偏波 (σ_{HH}^0)	HV 偏波 (σ_{HV}^0)	VV 偏波 (σ_{VV}^0)
	CP parameters	エントロピ (H)	アニソトロピ (A)	アルファ角 (α)
	VZ parameters	二回散乱 (Pd)	体積散乱 (Pv)	表面散乱 (Ps)
TerraSAR-X	後方散乱係数	HH 偏波 (σ_{HH}^0)	VV 偏波 (σ_{VV}^0)	
	CD parameters	エントロピ (H)	アルファ角 (α)	

方散乱係数のほかに、山谷ほか (2017b) において分類の精度向上に有効であった散乱モデル分解法である Van Zyl 法 (Van Zyl *et al.*, 2011) および固有値解析法である Cloude-Pottier 法 (Cloude and Pottier, 1996) による散乱成分も使用した。RADARSAT-2 データからは、以上の3種のデータセットの計9要素を5時期すべてのデータから算出し、計45の変数を分類に使用した。以降では、後方散乱係数、散乱モデル分解法、固有値解析法のデータセットを、それぞれ σ^0 , VZ, CP と表記する (表2)。

同様に、TerraSAR-X データからも各偏波の後方散乱係数 (σ_{HH}^0 , σ_{VV}^0) を式(2)から算出した。

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

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。このほか、固有値解析法である Cloude 法 (CD) (Cloude, 2007) の散乱成分 (エントロピ (H), アルファ角 (α)) を算出した。TerraSAR-X データからは、以上2種のデータセットの計4要素を5時期すべてのデータから算出し、合計20の変数を分類に使用した。

2.3 解析方法

本研究では、対象地域における圃場境界 GIS データを用い、圃場ポリゴン単位で作付作物の分類を行った。まず、道路や防風林などの混入および影響を防ぐため、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。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サイズがそれぞれ 8m, 2.75m であることを考慮し、バッファは10mに設定した。次に、算出したそれぞれの変数について、圃場ごとに平均値を抽出した。その後、分類処理のための教師データとして、Hartfield *et al.* (2013) に倣い20%の圃場を無作為に抽出し、残りの80%をテストデータとした。教師データと

表3 各圃場数の内訳

作物	圃場数		
	教師データ	テストデータ	合計
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

テストデータに用いた作物種別の圃場数を表3に示す。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。

分類には機械学習アルゴリズムであるランダムフォレスト (Breiman, 2001) を使用した。ランダムフォレストは、変数の一部を使用した決定木を複数作成し、それらの決定木から算出された結果の多数決により、最終的な分類を決定する手法である。本研究では、決定木の数を200本、1本の決定木を作成する上で使用する変数の数は Breiman (2001) に倣い、使用した全変数の平方根に設定して分類を行った。なお本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 (R Core Team, 2014) 及び、randomForest パッケージ (Liaw and Wiener, 2002) を活用した。

2.4 分類処理の流れ

本研究では、山谷ほか (2017b) で最も高い精度が得られた RADARSAT-2 のデータセット (σ^0 , CP, VZ) に、TerraSAR-X から算出されたデータセット (σ^0 , CD) のうちいずれか、または両方を併用することによって、分類精度の向上を図った。

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, OA) 及びカッパ係数 (κ) の2種類である。その他、作物ごとの分類精度の評価に、参照データの精度を示すプロデューサーズ精度

(PA) 及び分類結果の精度を示すユーザズ精度 (UA) の2種類を使用した。また、2つの手法間における κ 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つの分類結果に有意な差があると判断することができる。

3. 結果と考察

3.1 RADARSAT-2 データに TerraSAR-X の片方のデータセットを併用した分類

全ての分類結果から算出された OA と κ を図2に示す。図2において、RADARSAT-2 の3データセット (σ^0 , CP, VZ) を RS2 とし、TerraSAR-X の両方のデータセット (σ^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 (σ^0 : OA = 0.928, κ = 0.911, CD: OA = 0.929, κ = 0.912) は、

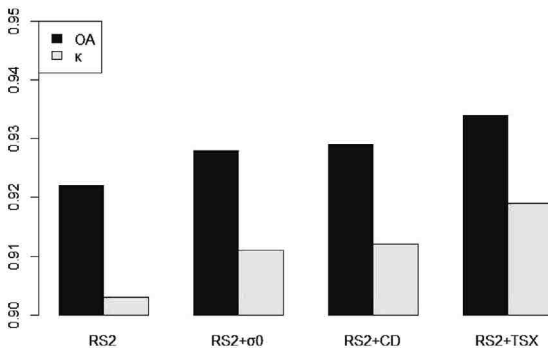


図2 各分類結果の全体精度とカッパ係数

表4 各分類結果で使用したデータセットの組み合わせ

	RADARSAT-2			TerraSAR-X	
	σ^0	CP	VZ	σ^0	CD
RS2のみ	○	○	○		
RS2+ σ^0	○	○	○	○	
RS2+CD	○	○	○		○
RS2+TSX	○	○	○	○	○

表5 分類結果間の検定統計量 Z

	RS2のみ	RS2+ σ^0	RS2+CD	RS2+TSX
RS2のみ				
RS2+ σ^0	1.100			
RS2+CD	1.275	0.175		
RS2+TSX	2.201	1.102	0.927	

表6 TerraSAR-X の1データセットのみを併用した作物別の分類結果

	RS2+ σ^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
Kappa		0.911		0.912

どちらも RADARSAT-2 のみでの分類結果 (OA = 0.922, κ = 0.919) より高い精度となっていた。しかし、これらの間には5%の有意水準で有意な差がなく (表4), 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結果を表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見られた。この理由として、小麦は生育ステージが他の作物より早く、また散乱形態も他の作物と大幅な差異があったためであると考えられる。一方、トウモロコシは豆類や馬鈴薯との間で誤分類が多く生じたため、群落における散乱の形態が近かったと示唆される。また、 σ^0 を併用した場合と CD を併用した場合では有意な差のない結果となったが、両者の差異として、前者は小麦で CD 併用より高い PA, UA が得られ、後者は馬鈴薯で σ^0 併用より高い PA, UA が得られるなど、異なる特徴があることも明らかとなった。

3.2 RADARSAT-2 データにすべての TerraSAR-X データセットを組み合わせた分類

TerraSAR-X から算出された2種類全てのデータセットを RADARSAT-2 に併用したところ、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、5%の有意水準で有意な差があった ($Z = 2.20 > 1.96$, 表4)。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て分類を

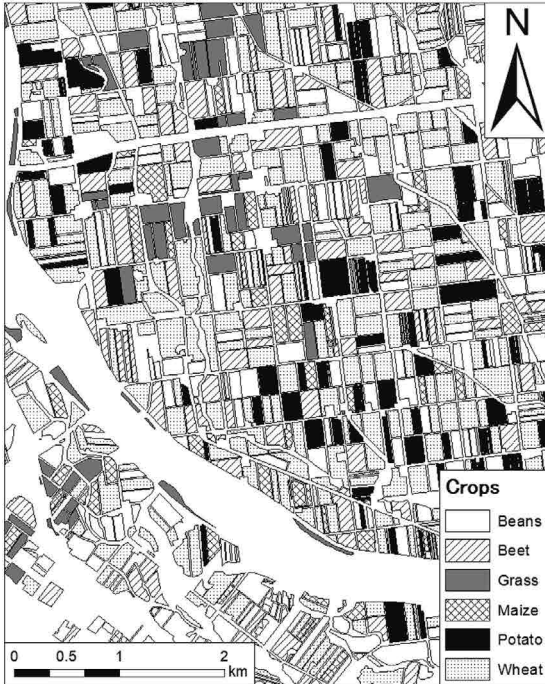


図3 RADARSAT-2 および TerraSAR-X のデータを組み合わせた分類マップ (一部)

表7 RADARSAT-2 および TerraSAR-X のデータを組み合わせた分類結果

		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	778	13	5	30	18	4	848	0.917
	甜菜	8	408	2	3	1	3	425	0.960
	牧草	9	2	491	15	9	4	530	0.926
	トウモロコシ	13	11	9	421	12	3	469	0.898
	馬鈴薯	16	7	5	28	551	7	614	0.897
	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		
									全体精度 0.934
									カッパ係数 0.919

行うことで、いずれかの衛星のみを使用した場合と比較して、高い精度での分類が行えることが明らかとなった。この理由として、双方から得られる情報が補完し合うことで精度が向上するためであると考えられる。最後に、この分類結果のマップを図3に示す。

TerraSAR-Xの2データセットを併用した分類結果の判別効率表を表7に示す。作物別に見た場合、TerraSAR-Xを1データセットのみを併用した場合には、それぞれの結果でPA, UAの低い作物が異なっ

ていた。しかし、両方のデータセットをRADARSAT-2に併用することで、いずれかで低かった作物の精度が補われ、全体的に高い精度へと繋がったことが明らかとなった。

4. 結 論

本研究では、Cバンドの4偏波SARデータおよびXバンドの2重偏波SARデータを各5時期ずつ利用し、機械学習アルゴリズムのランダムフォレストによる作付作物の分類を行った。その結果、Xバンドデータから算出されるデータセットの1つをCバンドデータに併用した場合、圃場の作付作物をCバンドの4偏波SARデータのみを用いた場合の0.92を上回る全体精度で分類可能であった。しかしこの結果は、Cバンドデータのみを使用した場合と精度に有意差はなかった。一方でXバンドのデータセットを2種類使用することで、全体精度は0.93を超えただけでなく、Cバンドのみを使用した場合と比較して有意な差のある結果が得られた。したがって、5時期のCバンド、Xバンド衛星データから複数のデータセットを算出することで、いずれか一方の衛星データを使用した場合よりも高い全体精度で分類可能であることが示された。この理由として、XバンドとCバンドの波長の違いから、植被層での透過の度合いが異なることが考えられる。Xバンドは波長が短く作物体を透過しにくいいため、作物体の表面付近での散乱で群落の表面構造に基づく情報を与えたと考えられる。一方Cバンドは、波長が長く比較的作物体を透過しやすいため、土壌表面や作物体の内部での散乱で群落の内部構造に基づく情報を与えたと考えられる。すなわち、透過の度合いが異なる2種類のバンドの衛星データを併用したことにより、更に高い精度での分類がなされたと結論付けられる。これまでXバンドやCバンドの衛星データを単独で利用することにより、高い精度での作付け作物分類が行われてきたが、分類において特に高い精度を目標とする場合に、本研究において実行した2種類のバンドを併用する手法の有効性が示された。

本研究では、2種類のバンドの衛星データを使用したことにより、その分データの購入におけるコストが高くなる課題がある。したがって、より少ない時期数で分類を行うことにより、少ないデータ数でも高い精度での分類を行える手法について検討する必要がある

る。また分類アルゴリズムに関しても、他の新しいアルゴリズムを適用した分類を行い、高精度で効率の良い手法を検討する予定である。

謝 辞

本研究を遂行するにあたり使用した圃場境界 GIS データは、十勝農業共済組合より提供して頂いた。記して謝意を表する次第である。

(受付日2018.1.9, 受理日2018.3.13)

引用文献

- 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.
- 国土地理院, 2016. 基盤地図情報ダウンロードサービス, Ibaraki, Japan.
<http://fgd.gsi.go.jp/download> (accessed 7 Mar. 2017).
- 山谷祐貴, 藺部 礼, 谷 宏, 王 秀峰, 小林伸行, 望月貫一郎, 2017a. TerraSAR-X データを使用したランダムフォレストによる作付作物の分類. 北海道大学農学部紀要, 34(2), pp.1-11.
- 山谷祐貴, 藺部 礼, 谷 宏, 王 秀峰, 小林伸行, 望月貫一郎, 野田 萌, 2017b. Cバンド SAR データを利用した機械学習アルゴリズムによる圃場の作物分類. 写真測量とリモートセンシング, 56(4), pp.143-148.
- 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.
- 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., Pottier, E., 1996. A review of target decomposition theorems in radar polarimetry. IEEE Transactions on Geoscience and Remote Sensing, 34(2), pp.498-518.
- Congalton, R.G. and K. Green, 2008. Assessing the accuracy of remotely sensed data: Principles and practices. CRC Press, Boca Raton, Florida, United States, pp.137.
- 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.
- Jiao, X.F., Kovacs, J.M., Shang, J.L., McNairn, H., Walters, D., Ma, B.L., 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.
- Liaw, A., M. Wiener, 2002. Classification and Regression by Random Forest. R News 2, pp.18-22.
- NGA GEOnet Names Server, 2013. Earth Gravitational Model 2008 (EGM2008), United States.
<http://earth-info.nga.mil/GandG/wgs84/gravitymod/egm2008> (accessed 7 Mar. 2017).
- R Core Team, 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/> (accessed 29 Jan. 2017)
- Sonobe R., Tani H., Wang X., Kobayashi N., Shimamura H., 2014. Random forest classification of crop type using multi-temporal TerraSAR-X dual-polarimetric data. Remote Sensing Letters, 5(2), pp.157-164.
- Van Zyl, J.J., Aari, M., 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.
- Xu, J., Li, Z., Tian, B.S., Huang, L., Chen, Q., 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.