Int J Appl Earth Obs Geoinformation 67 (2018) 114–122

Contents lists available at ScienceDirect

Int J Appl Earth Obs Geoinformation

journal homepage: www.elsevier.com/locate/jag

# Coefficient of variation for use in crop area classification across multiple climates

Tracy Whelen, Paul Siqueira*

*Department of Electrical and Computer Engineering, 113 Knowles Engineering Building, 151 Holdsworth Way, University of Massachusetts, Amherst, MA, 01003, United States*

ARTICLE INFO

*Keywords:*
SAR
Backscatter
Coefficient of variation
Time series
Agricultural land cover classification

ABSTRACT

In this study, the coefficient of variation (CV) is introduced as a unitless statistical measurement for the classification of croplands using synthetic aperture radar (SAR) data. As a measurement of change, the CV is able to capture changing backscatter responses caused by cycles of planting, growing, and harvesting, and thus is able to differentiate these areas from a more static forest or urban area. Pixels with CV values above a given threshold are classified as crops, and below the threshold are non-crops. This paper uses cross-polarized L-band SAR data from the ALOS PALSAR satellite to classify eleven regions across the United States, covering a wide range of major crops and climates. Two separate sets of classification were done, with the first targeting the optimum classification thresholds for each dataset, and the second using a generalized threshold for all datasets to simulate a large-scale operationalized situation. Overall accuracies for the first phase of classification ranged from 66%–81%, and 62%–84% for the second phase. Visual inspection of the results shows numerous possibilities for improving the classifications while still using the same classification method, including increasing the number and temporal frequency of input images in order to better capture phenological events and mitigate the effects of major precipitation events, as well as more accurate ground truth data. These improvements would make the CV method a viable tool for monitoring agriculture throughout the year on a global scale.

## 1. Introduction

As the earth is called upon to feed increasing numbers of people with a limited supply of arable land and water, accurate, up-to-date monitoring of global agriculture becomes an increasingly important part of world stability for reasons of food security, economic stability, climate change, and environmental degradation (Becker-Reshef et al., 2010; Jayne and Rashid, 2010; and Scherr and Sthapit, 2009). Remote sensing provides a way to monitor agriculture at a global scale, with reasonable time and manpower requirements while also providing a uniform system of measurement. Due to the dynamic and complex nature of agricultural landscapes, with hundreds of crop types growing on fields ranging from tenths to hundreds of hectares planted in numerous climatic conditions, remotely sensed measurements of agricultural land area vary drastically across the globe in their precision, accuracy, and timeliness, with global operationalization still proving elusive (Waldner et al., 2015). Most global land cover datasets include agricultural lands as part of mosaic or mixed classes, variably including pasture, which makes them challenging to use for agricultural applications (Bartholomé and Belward, 2005; Bontemps et al., 2011; Friedl et al., 2010). The few with dedicated agriculture classes struggle with accuracy (Gong et al., 2013) or otherwise highlight the uncertainty and challenges in estimating global cropland extent (Biradar et al., 2009; Pittman et al., 2010; Ramankutty et al., 2008; Yu et al., 2013). These discrepancies stem from a number of issues including the availability of cloud-free images at the desired spatial and temporal resolution of all the necessary regions of the globe, and the availability of detailed ground truth data for training classification algorithms (Matton et al., 2015 and Whitcraft et al., 2015).

As an active remote sensing system, synthetic aperture radar (SAR) can help mitigate some of the challenges of optical imagery, as the data are mostly independent of solar and atmospheric conditions, allowing for reliable data collection in areas of frequent cloud cover. Increasing numbers of SAR-based agricultural land cover classifications have been made using SAR as a standalone source, as well as in combination with optical data. Most of these works have focused on differentiating individual crops; however they primarily have been non-operational projects focused on small regions (Champagne et al., 2014; Haldar et al., 2012; Skriver et al., 2011). Operational projects primarily consist of Agriculture and Agri-Food Canada's annual end-of-season crop map

* Corresponding author.
*E-mail address:* siqueira@umass.edu (P. Siqueira).

https://doi.org/10.1016/j.jag.2017.12.014
Received 7 August 2017; Received in revised form 21 November 2017; Accepted 26 December 2017

(Fisette et al., 2014) and a few rice monitoring projects in Asia (Chakraborty et al., 2006 and Nelson et al., 2014). Crop/non-crop classifications using SAR are not operationalized, and have primarily been related to monitoring agricultural land abandonment, but again over relatively limited regions and crops (Stefanski et al., 2014, and Yusoff et al., 2017).

Past work has found that L- and C-band are the most effective wavelengths for agricultural applications due to their sensitivity to both fine-scale structural characteristics of different crop types and growth stages, as well as to soil moisture and other soil characteristics (Ferrazzoli et al., 1997 and McNairn and Shang, 2016). Multi-temporal datasets have been found to be more effective at agricultural land cover classification due to their ability to capture agricultural phenology and related growth patterns (Blaes et al., 2005; Deschamps et al., 2012; Jiao et al., 2014). In order to capitalize on the changing radar signature over time exhibited by agricultural areas, as opposed to the relatively static signature of a forest or urban area, we introduce a statistical measurement known as the coefficient of variation (CV) to measure the variation in backscatter response over time for a single location. More commonly used to measure spatial variation, here the CV is used to measure temporal variation. It focuses on the overall variability and generally dynamic nature of crop planting, growing, and harvesting cycles, not specific phenological features of individual species. The significant changes in ground cover, physical structure, and backscatter of an agricultural area is expected to produce a wider range of backscatter values over time, and thus a higher CV (Cihlar et al., 1992). Once the CV has been computed for an entire region, a threshold can be used to classify pixels as a crop or non-crop, with crops having the higher range of CV values. Crop/non-crop classifications are valuable not only as a predecessor to classifications differentiating between individual crops, but also for what they can tell about land use patterns, such as continued farming in a conflict zone, or spotting new agricultural lands in a previously forested region. Other types of thresholds have been previously used for SAR-based classifications, as an example, for rice (Bouvet and Le Toan, 2011).

The methodology presented here is chosen because of its ability to be more statistically traceable than common methods of classifying a multi-temporal range of images, such as decision trees or the maximum likelihood classifier. As a simple hypothesis test using a statistics-based input layer, the CV classification allows one to estimate the effect on the classification error if a different threshold and/or number of input images were used. This characteristic makes this classification method useful both when designing data collection patterns for upcoming satellite missions, and for data users trying to select the minimum number of already collected images needed to produce a classification with a specified target accuracy. For crop/non-crop classifications, simple algorithms that do not depend on extensive training data and work in a wide range of climactic conditions have value for providing global measurements of agricultural area (Matton et al., 2015 and Waldner et al., 2015). The CV algorithm fits these preferred characteristics of being computationally simple, requiring minimal training data, and using consistent methodology across the globe.

An important application of this work is connected to NISAR, a joint project between NASA and the Indian Space Research Organization (ISRO). This mission, planned for launch in 2021 will be primarily collecting L-band HH/HV data over most land surfaces worldwide with a 12 day repeat cycle (Rosen et al., 2015). Its consistent, freely available (NASA, 2012) global coverage will provide a unique opportunity to investigate applications using longer and more frequent time series (some 30 images per year) than what has traditionally been feasible due to data cost and coverage (Fisette et al., 2014). Further investigation of applications techniques is encouraged by the NISAR project's Level 1 Science Requirements, which include the requirement that the satellite be able to seasonally classify global croplands at 80% or better at the one hectare scale (Sanchez, 2014). This paper not only tests the potential for the CV method in preparation for the NISAR satellite, it also shows how NISAR data could be used to improve global agricultural monitoring.

## 2. Data

Preliminary research has shown that cross-polarized L-band imagery is more effective at separating crop and non-crop regions using the CV classification method than C-band or co-polarized L-band observations (Siqueira, 2016). Because of these findings, cross-polarized (HV) imagery from the Japanese L-band ALOS satellite was used for this research, as it is one of the few available options for repeat-pass L-band SAR data. In addition, ALOS data spans a wide range of ecosystems and crop types, including coverage across all of the United States, and has previously been used for agricultural land cover classification (Haldar et al., 2012; McNairn et al., 2009 and Yusoff et al., 2017). Terrain corrected and ground projected images are available from the Alaska Satellite Facility (Alaska Satellite Facility Engineering Group, 2015). All images were acquired in fine beam dual mode (FBD), ascending orbits, with consistent viewing geometry including an off-nadir viewing angle of 34.3° and incidence range of 36.6°–40.9° (Rosenqvist et al., 2014). ALOS-1 had a very consistent orbit track, and hence incidence angles did not change over time. The consistent orbit also meant that data was repeatedly captured over the same strips of land, with significant overlap between neighboring strips. Using imagery from ASF provided consistent calibration and preprocessing, producing multilooked 30 m pixels with backscatter recorded in terms of gamma nought, which is a terrain-corrected version of the normalized radar cross section (Small, 2011). ALOS operated on a 46 day repeat cycle; however coverage locations varied between cycles. Each of the test locations has between six and fourteen images taken between 2007 and 2010 during the spring, summer, and fall months, with analysis using all dates that provided a complete image strip for the given location.

Ground truth information came from an agricultural land cover classification product known as the Cropland Data Layer (CDL), which is released annually by the United States Department of Agriculture National Agricultural Statistics Service (USDA NASS) (Boryan et al., 2011). The CDL was chosen as a relatively high resolution land cover database produced for a large region under consistent methodology; however its use did limit the possible study areas to the continental United States. Production of the CDL combines optical data primarily from Landsat 5 TM, Landsat 7 ETM+, and RESOURCESAT-1 AWiFS sensors with high quality ground truth data as input for sophisticated decision-tree classification software. The resulting product has an overall accuracy of around 80% for all agricultural land covers, with major crops such as corn, soybeans, wheat, cotton, and rice frequently having accuracies over 90%. While not 100% accurate, the CDL is a nationally available source of agriculturally-focused land cover data that is a more consistent source of data than if separate ground truth datasets had been used for the different regions of the country. At a national scale, classifying individual crops, with 30 m or 56 m pixels (depending on the year) it is one of the best publically available agricultural land cover databases available in the world, and as such has been used as training data for attempts at satellite-based global crop mapping (Matton et al., 2015; Pittman et al., 2010; Waldner et al., 2015).

Because this paper focused on a crop/non-crop classification rather than classifying individual crop types, the full CDL information was simplified into a multi-year crop/non-crop layer. Since the CDL is known to have difficulty classifying fields containing multiple crops in a single year, either due to interplanting of rows or double cropping that alternates summer and winter crops, a crop/non-crop version should reduce the influence of these types of errors (Boryan et al., 2011). It should also mitigate errors related to confusion between similar crops, such as oats or barley being misclassified as wheat. While the ALOS imagery spans 2007–2010 growing seasons, and production of the CDL began in the late 1990's with a few key producing states, the CDL is

only available nationwide from 2008 on (Boryan et al., 2011). The 2008–2010 layers were used to define a crop/non-crop mask where a pixel was classified as a crop if it had been classified as some sort of field crop for at least two of the three input years. The term "field crops" is used here to refer to all annual crops such as grains and vegetables, as well as perennials such as alfalfa or sugarcane that are harvested at least once a year and replanted every few years. In addition fallow fields were included, as some types of crops such as winter wheat are fallow roughly every other year. In contrast to the cultivation layer released as part of the CDL, woody long-standing perennials like grapes and fruit and nut trees were not included.

The choice of dataset locations was a balance between existing growing locations of key crops, climate, and available coverage of ALOS imagery. Staple crops of interest to the NISAR mission include barley, canola, corn, potato, rice, rye, soybean, sugar beet, sugarcane, sunflower, and wheat, most of which have significant acreage in multiple areas around the United States (Siqueira and Saatchi, 2015). Each dataset includes two or three of these staple crops as major crops, often an additional non-edible major crop like cotton or alfalfa, and a smattering of minor crops such as vegetables and oil seeds. While individual crops have certain climatic and locational restraints as to where they can be successfully grown, agriculture as a whole spans a wide variety of ecosystems. To capture some of this site diversity, the decision was made to intentionally choose sites across a wide range of climates. The USDA Plant Hardiness Zones (Daly et al., 2012) and a recent version of the Köppen-Geiger Climate Classification (Peel et al., 2007) were used to guide this portion of dataset site choice. The final definitive factor when choosing a dataset was the availability of ALOS imagery. A given site needed to have at least seven dates available for every image location in the dataset. Seven was chosen as a lower threshold with the goal of having multiple dates of images from most years of data collection. Two of the strips (one each in California and Texas) ended up with only six images in the final processing due to some differing image dates between the strips. The combination of these three constraints – existing staple crops, climate, and image availability – guided the selection of datasets to eleven sites ranging from Eastern Washington to Southern Georgia (Table 1 and Fig. 1).

## 3. Methods

The CV is a unitless statistical measurement of variation, defined as

$$CV = \frac{std(\sigma)}{mean(\sigma)} \quad (1)$$

where in this case the distribution refers to the set of backscatter values for one pixel location over the time series of images (Kendall and Stuart, 1969). The distribution of backscatter values captures the seasonal variation in crops as opposed to the relatively static response of non-agricultural land. Here the CV is used in a novel fashion to measure temporal variation such that it can be used for crop/non-crop classification, as opposed to the more common use of the CV to measure spatial variation.

Creation of the full CV layer for each dataset involved three steps of mosaicking and calculation. The ALOS-1 orbit pattern was such that adjacent strips of data were collected on different days, and two neighboring orbit strips might have differing numbers of total images (Fig. 1). Because of this, the individual images were first mosaicked into orbit strips, then the image strips were combined to create a CV strip for each orbit location, and finally the CV strips were mosaicked together into a full CV mosaic of the entire dataset (Fig. 2). Along the way pixels were aligned, a mask of all water-covered areas (lakes, rivers, etc.) based on the CDL data was applied, and edge pixels covered by only a subset of images were excluded.

Two phases of classification were employed to classify each dataset's CV layer into crop and non-crop, both of which are visualized in Fig. 3. The first phase involved finding the optimal classification threshold of each individual dataset using two different techniques. One technique was using a receiver operating characteristic (ROC) curve and the other was the best separation of histograms. The ROC curve of a dataset is created by graphing the true positive rate (TPR) against the false positive rate (FPR) at a series of thresholds. The ROC curve for a completely random test that does not discriminate between the two populations would have a straight diagonal line where the TPR and FPR are identical. The farther away an ROC curve is from this line, the better it can differentiate between the two populations of data (Fawcett, 2006). ROC curves were created for each dataset by classifying the CV layer and comparing it against the CDL-derived crop/non-crop mask, with the series of thresholds ranging from 0 to 1 with a step size of 0.01. The threshold where the ROC curve had the greatest separation from the line of no discrimination was designated the best classification threshold using the ROC method. The second technique for finding a threshold of best classification was done by creating histograms that graphed the number of pixels classified by the CDL-derived mask as

**Table 1**
Introduction to the dataset site locations giving name, percentage covered by crops according to the CDL, size, major land covers, and Köppen-Geiger climate type. Size (scenes) refers to how many row/path locations were used, and is referring to the same amount of land given in Size ($km^2$). To see the layout of the scenes at a given site or the number of dates used over any given scene location refer to Fig. 1.

| Map Key | Dataset | % CDL crops | Size (Scenes) | Size ($km^2$) | Major Agricultural Land Covers | Other Main Land Covers | Climate Classification[a] |
|---|---|---|---|---|---|---|---|
| 1 | Eastern Washington | 25.5 | 25 | 63009 | Winter Wheat, Alfalfa, Fallow, Potatoes | Shrubland, Grassland, Forest | BSk, Csb |
| 2 | Central California | 28.3 | 20 | 62469 | Alfalfa, Cotton, Corn, Winter Wheat, Rice | Grasslands, Developed, Shrubland | BSk, Csa |
| 3 | Idaho West | 18.5 | 13 | 34282 | Alfalfa, Barley, Sugarbeet, Potatoes, Corn, Winter Wheat | Shrubland, Grassland, Forest, Barren | BSk |
| 4 | Idaho East | 25.8 | 5 | 15138 | Barley, Alfalfa, Potatoes, Spring Wheat | Shrubland, Forest | Dfb, BSk |
| 5 | North Dakota | 33.6 | 25 | 60723 | Spring Wheat, Durum Wheat, Peas, Lentils | Grasslands, Forest | Dfb, BSk |
| 6 | Minnesota | 18.2 | 15 | 37543 | Corn, Soybeans, Alfalfa | Grassland, Forest, Developed, Wetlands | Dfb |
| 7 | West Kansas | 53.8 | 30 | 101958 | Corn, Winter Wheat, Fallow, Sorghum, Alfalfa | Grassland | BSk, Cfa |
| 8 | East Kansas | 39.1 | 18 | 50990 | Soybeans, Winter Wheat, Corn, Sorghum, Alfalfa | Grassland | Cfa, Dfa |
| 9 | South Texas | 16.8 | 13 | 36886 | Sorghum, Cotton | Shrubland, Grassland, Developed | Cfa |
| 10 | Mississippi Memphis | 39.3 | 40 | 111976 | Rice, Soybeans, Cotton, Corn | Forest, Grassland, Wetlands | Cfa |
| 11 | Georgia | 15.3 | 9 | 27712 | Cotton, Corn, Peanuts | Forest, Grassland, Wetlands | Cfa |

[a] See Peel et al., 2007 for descriptions of the different climate classifications.

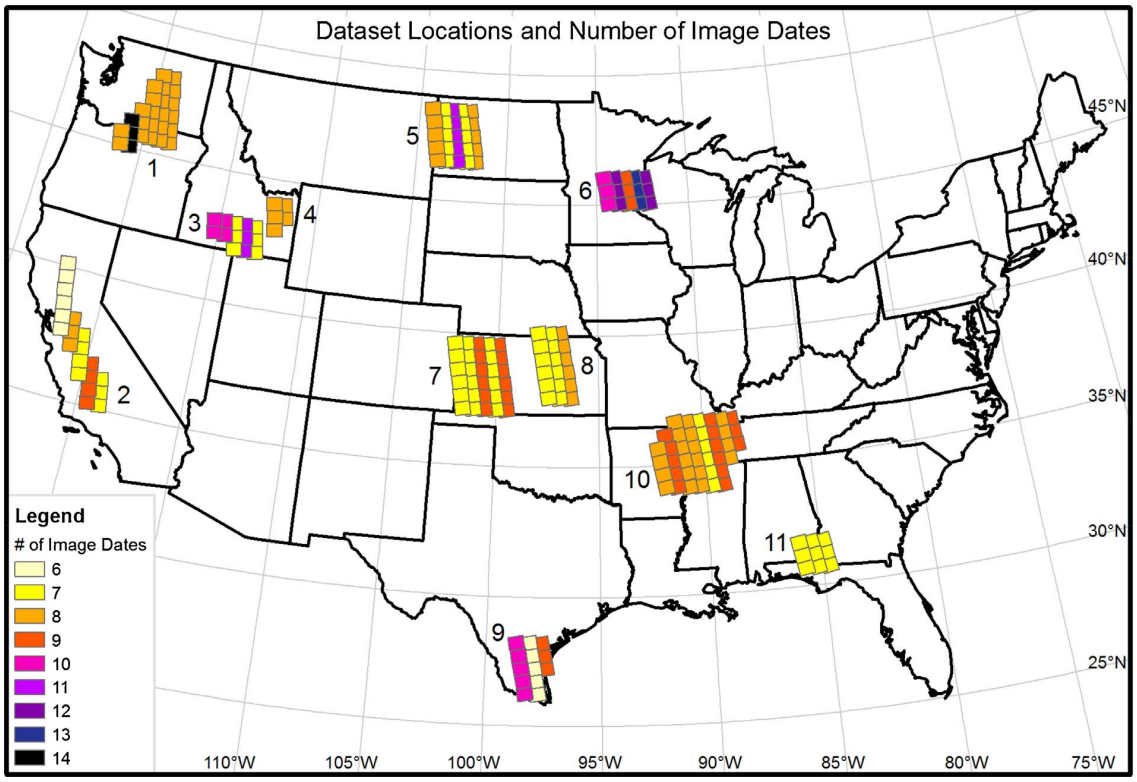


**Fig. 1.** Diagram showing locations of the datasets used in the analysis. Numbers correspond to the entries in Table 1, which give the site names and further details about each site. The colors correspond to the number of images used over a given scene location.

crop or non-crop against the range of CV values. The point of least overlap between the two histograms was used as the histogram method threshold. The ROC curves and histograms for each dataset can be seen in the Supplemental Materials that accompany the online version.

The second phase involved using a single threshold based on the results of the first phase to classify all of the CV layers again. The goal of this second phase of classification was to see how much the classifications changed when the threshold was not fine-tuned to fit each individual dataset, much as might be required in an operationalized setting. The threshold used in phase two was chosen from the range of best thresholds produced in phase one, however the two phases of classification were otherwise separate and not sequential. For this project a CV of 0.5 was used as the classification threshold for phase two, given that the thresholds produced in phase one ranged from 0.35 to 0.60.

Classification accuracy was assessed by calculating the overall, user's, and producer's accuracies for the classifications based on all three thresholds for each dataset (ROC, histogram, and phase two). To compute the accuracies, all pixels of the classification were compared against the CDL-derived multi –year crop mask that is described in the Data section. Full details of the accuracy results for each dataset can be found in the Supplemental Materials.

## 4. Results

Phase one of the classification step produced thresholds ranging from 0.35 to 0.60. The two methods of threshold development, ROC curves and histogram separation, produced very similar thresholds. For all eleven datasets the ROC-based threshold was equal to or up to 0.03 higher than the histogram-based threshold, with the exact numbers shown in Table 2. The overall classification averages in phase one ranged from 66% to 81%. When looking at the user's and producer's accuracies (see Supplementary Materials) the non-crop land covers tended to have slightly higher accuracies than the crop areas. This is related to the fact that all but one of the datasets had more non-crop pixels than crop pixels. A different threshold could have correctly classified more crop pixels, but the drop in accuracy for the non-crop pixels would have worsened the overall classification. In phase two of classification, using a threshold of 0.5 for all datasets, the overall accuracies ranged from 62% to 84%. Increases and decreases in accuracy were split equally, with five datasets increasing, five decreasing, and one that had already had an ROC-based threshold of 0.5 staying constant. The highest overall accuracies were when the two types of error were balanced, with roughly equal numbers of false positives and false negatives. Moving closer to or farther away from this equilibrium defined which datasets had their accuracies respectively rise or fall during phase two of classification. The datasets with phase one thresholds farthest away from 0.5 had the greatest change of accuracy between the two phases.

## 5. Discussion

The intended eventual use for the CV algorithm is to produce a global crop/non-crop product to improve on the existing products discussed in the introduction. With these benchmarks in mind, the results from this work show a variation of classification successes and challenges posed by heterogeneity. Overall accuracies in this project were slightly lower than the 71%–99% found by Matton et al. (2015), and comparable to those found by Pittman et al. (2010). Other projects such as Yu et al. (2013) and Ramankutty et al. (2008) are harder to compare against since they measured accuracy by assessing acreage per agricultural census unit (country, provinces, etc.) instead of per pixel or other smaller validation unit.

While many fields were correctly classified by the CV method, a

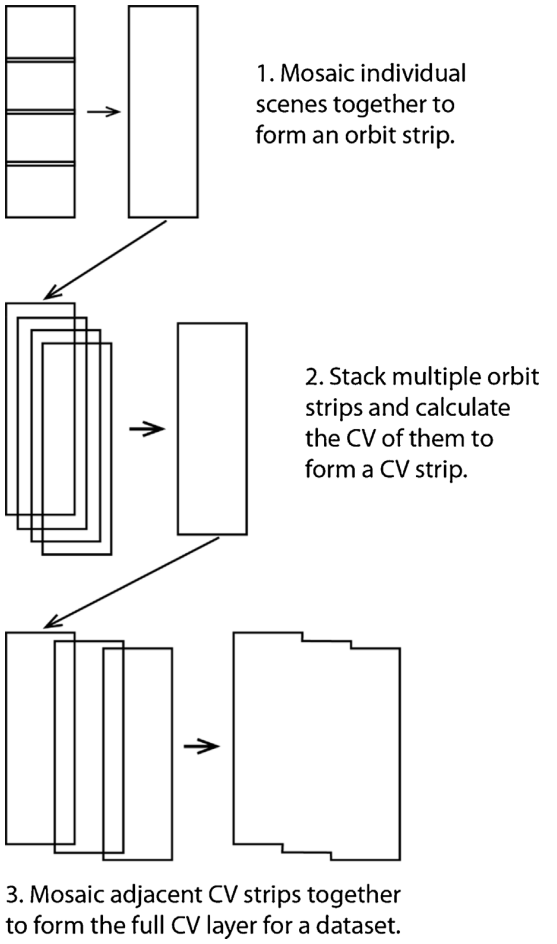


**Fig. 2.** Diagram showing processing steps used to create the full CV layer for each dataset.

close examination of each of the datasets revealed multiple potential causes for the misclassification of pixels, and potential solutions related to both input data and methods. The major sources for error were the effect of recent rain events and challenges classifying agricultural land covers with minimal variation like alfalfa and fallow fields, both of which are related to the overarching problem of suboptimal image timing. Seven of the datasets (Eastern Washington, Mississippi Memphis, South Texas, West Kansas, North Dakota, Idaho East, and Idaho West) had a distinct line of errors along a strip edge (example in Fig. 4). At all seven of these sites there was a rainstorm over at least one of the images in that strip (NOAA/NWS, 2017) less than 48 h before image capture. Seven datasets (Central California, Eastern Washington, East Kansas, West Kansas, North Dakota, Idaho East, and Idaho West) had significant errors for minimally changing crops such as alfalfa or winter wheat/fallow crop rotations. Both of these issues highlight the more widespread issue of suboptimal image timing that would be improved by using more frequent images captured during a single season, such as will be provided by the NISAR mission. Due to the number of potential error sources and large area under analysis it is not possible to definitively determine specific source(s) of error for each misclassified region. An example of both correct classification and a few different types of error can be seen in Fig. 4, with additional details following.

When viewed manually, individual fields tended to have consistent classification, where most fields had well over 80% of their pixels classified either correctly as a crop or incorrectly as a non-crop (Fig. 4). Uncommon were fields whose pixels were an even mixture of crop and non-crop. Completely misclassified fields point to issues like poor ground truth, limited structural variation over time for an entire crop (like alfalfa), or abnormal growth patterns for a single field, all of which could potentially be solved with better input data and additional processing. The relatively few fields with mixed results indicate that the CV algorithm is a good discriminator for agricultural disturbance because it is clear that even though it is a pixel-level classifier, large-contiguous areas that are clearly field crops appear to be correctly classified. In addition, most of the major crops, including corn, soybeans, rice, and cotton classified well. The two exceptions were wheat and alfalfa, which are discussed in detail below with the other main sources of error. Because individual fields tended to be classified as almost all crop or almost all non-crop, this suggests that some type of spatial smoothing or a per-field classification could improve the overall results, as has been suggested by previous work (Fisette et al., 2014; Ma et al., 2014; Tso and Mather, 1999). Spatial smoothing could be applied to the original images by further multilooking or speckle filtering, or it could be applied at later stages with spline functions and increased pixel size. These techniques, all valid, were considered outside of the scope for this work. Including a classification step that reclassifies lone pixels that are completely surrounded by the opposite classification, setting a minimum field size for crop areas, or classifying entirely at the field level would be other ways to handle speckle-type misclassification.

Further analysis of the errors showed that false positives were in areas with major land use change, some grass or pasture areas, significant terrain, and with extra moisture. Areas with major land use change were primarily a few places in California, which shifted from grassland or fallow fields to vineyards or fruit and nut trees. Any of these land covers on their own would be clearly classified as a non-crop, but the shift mid-collection period made the CV layer appear similar to that of a crop, thus confusing the classification. Using imagery spanning only a single year would limit this type of error. The false positives in grass or pasture areas were particularly apparent in Georgia. When examined in Google Street View these regions appeared to be active agriculture areas, with activity such as growing straw or hay commercially, suggesting some sort of growth and harvest cycle is taking place. This is a harder error to solve, requiring a combination of more precise ground truth data and/or a better definition for what counts as a crop. The only area with significant topography was Eastern Washington, which caused some terrain-related false positives in regions where the DEM used for terrain correction differed from the topography of the actual terrain. Improvement of the DEM in these regions is expected to improve the types of errors found in these regions. False positives related to extra moisture appeared in two different forms. In some areas such as the South Texas dataset, the moisture could be seen in small depressions scattered across the region that potentially had extra vegetation growing around them at wetter times of the year. In other regions, like in western Idaho, the moisture was less visible in the vegetative ground cover, but appeared very clearly in images after recent precipitation. The authors surmise that this might be due to differing soil types, such as a patch of clay soils that retains moisture longer. Moisture is also potentially a factor in strip boundary errors, as discussed below.

False negatives had a noticeable correlation with agricultural land cover types with minimal change, wheat, and poor timing between the image dates and the crop calendars of certain species. Some land covers such as alfalfa and fallow fields are of an agricultural nature, and yet can have minimal variation over time, especially depending on the image collection dates. While not absolute, the CV method proved to have lower accuracy when classifying these types of fields. There were however multiple instances of two nearby fields growing alfalfa the entire time where one field was correctly classified and the other field was misclassified (as a whole). It is expected that in these situations, the limited variation and multiple harvests per year for alfalfa make the classification accuracy particularly sensitive to image timing as it relates to harvesting. Using a dataset with more frequent images would mitigate this sensitivity by increasing the potential for all fields to have

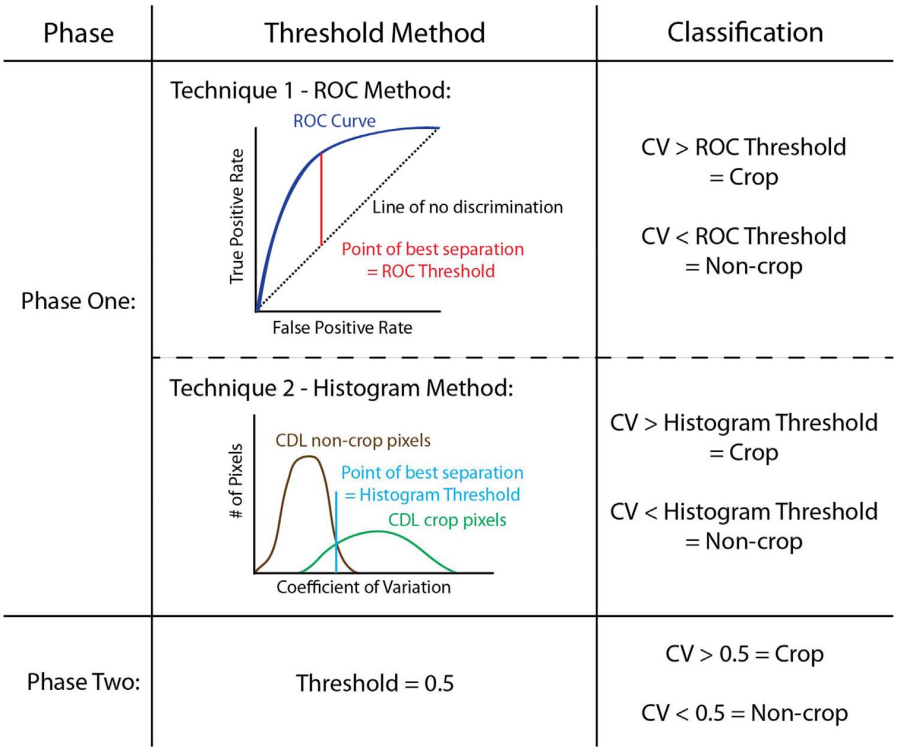


**Fig. 3.** Comparison of classification methods. The diagram shows the different phases of classification undertaken in this analysis. Phase one involved using an ROC curve and histograms to define thresholds for each individual dataset, and then phase two used a single threshold for all datasets that was determined based on the thresholds found in phase one.

**Table 2**
Classification accuracy results, showing the overall accuracy and threshold for all methods of classification for each of the eleven data sites.

| Dataset | Histogram Threshold | Histogram Overall Accuracy | ROC Threshold | ROC Overall Accuracy | 0.5 Threshold Overall Accuracy |
|---|---|---|---|---|---|
| Central California | 0.44 | 0.740 | 0.46 | 0.753 | 0.769 |
| Eastern Washington | 0.37 | 0.725 | 0.40 | 0.753 | 0.798 |
| Idaho West | 0.58 | 0.744 | 0.60 | 0.755 | 0.668 |
| Idaho East | 0.49 | 0.737 | 0.51 | 0.754 | 0.746 |
| North Dakota | 0.58 | 0.662 | 0.58 | 0.663 | 0.628 |
| Minnesota | 0.52 | 0.777 | 0.54 | 0.793 | 0.759 |
| West Kansas | 0.49 | 0.714 | 0.50 | 0.713 | 0.713 |
| East Kansas | 0.51 | 0.731 | 0.52 | 0.734 | 0.727 |
| South Texas | 0.45 | 0.813 | 0.45 | 0.811 | 0.848 |
| Mississippi Memphis | 0.41 | 0.811 | 0.42 | 0.815 | 0.820 |
| Georgia | 0.35 | 0.710 | 0.36 | 0.728 | 0.822 |

an image just after harvest. In addition, Hong et al. (2014) notes that in optical-SAR classifications involving alfalfa early season images are particularly important. Similarly, in some areas, such as Eastern Washington, standard growing practice is to alternate a field between growing winter wheat and leaving it fallow every other year. Given the minimal number of images that were available from any single year used in this research, this meant that there were difficulties in classifying wheat in some regions, as it appeared too much like an unchanging fallow field. This would be solved by using more frequent images, or images all from one growing season and then excluding fallow fields from the crop category. There is the additional potential that some feature about how wheat grows, in particular durum wheat, makes it harder to identify using the CV method. In North Dakota there were large sections of false negatives in areas that grew durum wheat, which in off years grew lentils or peas instead of being left fallow. Another possibility is that wheat's structure and/or growing patterns are of average difficulty to classify using SAR (suggested by Hoekman and Vissers (2003)), but the images captured for this region lined up poorly with the crop calendar for wheat and lentils. The NISAR mission's 12 day revisit time, and year-around consistent observation pattern will provide the more frequent imagery across the full growing season that is needed to improve classification in all of these cases.

Sources of error that resulted in both false positives and false negatives were broader issues that to some extent underlie this entire work. The first of these sources of error was misclassification in the CDL. While the CDL is the most accurate large scale agricultural land cover dataset over the entire United States, and by far the best available data source for the size and historical timing of our project, it is not 100% accurate. For major crops it is usually about 85%–95% accurate; however with minor crops it can have a lot more difficulty (Boryan et al., 2011). Much of this type of error is of little consequence in this work, since the multiyear CDL-derived cultivation layer is not affected by a crop being misidentified as some other type of crop. The problem with misidentification in the CDL comes when the classification changes the designation of crop or non-crop, either at the edges of a field or for a larger portion of a field. For example in parts of the Minnesota dataset the CDL had difficulty in distinguishing small fields of alfalfa and corn set amongst pasture, leading to some areas the CDL defined as pasture to be classified as crops. The cause of error becomes increasingly clear when comparing the field boundaries visible in the CV layer and Google Earth with the original CDL layers. In general, the potential effect of the occasional outlying misclassified pixel is negated by using the multi-year cultivation mask, which spatially smooths the data as compared to any single year. Using more accurate ground truth for future projects would reduce this source of error and provide better classification accuracy.

The other source of error that contributed to both types of error was when one strip had noticeably higher or lower values than its neighbor,

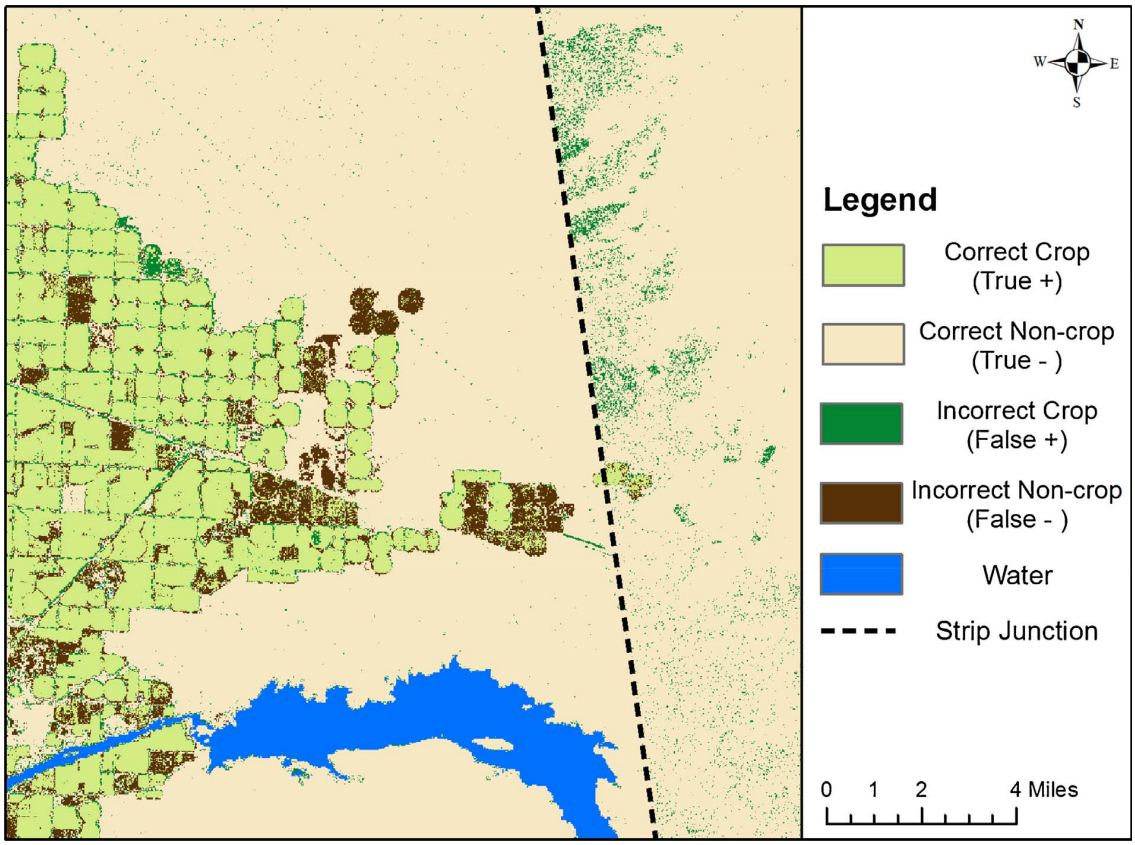


**Fig. 4.** CV classification example. In this portion of the Idaho West dataset you can see both correctly classified crop and non-crop regions, as well as a few different types of errors. The fields show how misclassification is mostly at the field scale, so that fields are mostly correctly or incorrectly classified, rather than having large amounts of speckle in all fields. The misclassified fields in this region are mostly areas that were alfalfa throughout image collection. They show the challenges in classifying crops with limited change in vegetation. On the right hand side of the image a distinct line of increase in false positives is due to the boundary between two strips of images.

causing a distinct line of increased misclassification at the boundary between the strips. This is potentially due to SAR viewing geometry, as well as the varying number of images used to make the CV layer of each strip. At the far edges of the swath backscatter values tend to be slightly different from the center of the image due to the difference in incidence angle, in this case by only a few degrees. When two adjacent strips overlap this becomes more apparent as the edge of one strip lies near the center of the neighboring strip. Another possible cause of the edge of swath errors is the differing dates and number of images used in the strip CV layers. Due to observing strategies and data availability, the images in neighboring strips were not collected on the same day, and not every strip had the same number of images. Adjoining images taken on different days may capture a precipitation event in one image but not the other image, despite the precipitation falling across both images. Increasing the number of input images, and making sure each strip has the same number of images would lessen the impact of any one precipitation event. The differences in CV values related to dates and number of images can become noticeable when strips are viewed side by side (Fig. 4). The portion of Idaho West in Fig. 4 and six other datasets (Eastern Washington, Mississippi Memphis, South Texas, West Kansas, North Dakota, and Idaho East) with similar distinct strip boundary lines had rain on or shortly before the capture of several of their images (NOAA/NWS, 2017). This type of moisture effect has been noticed as an issue in previous works involving mosaicking ALOS imagery (Lucas et al., 2010), but would be less of an issue with NISAR.

Precipitation and other seasonal climate differences have the potential to affect CV values between different regions. Uniform precipitation across the timespan of images would result in a smaller range of backscatter values, and thus a smaller CV. In terms of climate, this means that a region with greater variation in precipitation across the year, such as alternating between monsoon and dry seasons, is expected to have higher CV values than a region that receives consistent rainfall across the year. In this case, consistent precipitation could mean either high or low amounts of precipitation, given that the CV reflects change rather than absolute backscatter values. One of the goals of choosing datasets from around the United States in varying climates was to try to test these potential climate influences, either directly from the precipitation or the varying surrounding land covers that are related to precipitation and climate, such as grassland or forests. While single extreme precipitation events have been noticed to have a local effect on values, so far no conclusive patterns have been seen for broader climate effects. This is likely due to the datasets being comprised of a limited number of images collected across multiple years, as opposed to a dataset comprised of images collected every week or two for a single calendar year. More frequent data that is less susceptible to precipitation events, such as what will be provided by NISAR will provide a better opportunity to more fully test this hypothesis.

## 6. Conclusion

This paper presents the coefficient of variation as a unitless measurement used to make a binary crop/non-crop classification. This method is tested using L-band ALOS satellite images from across the United States taken over a four year period. Overall, using the approximately 80% accurate CDL as ground truth with a classification method that with perfect ground truth would correctly classify 80% of

the time, the expected combined accuracy would be around 64%, or 0.8 * 0.8. Given that all eleven datasets had noticeably higher accuracies than this, it is reasonable to surmise that if used with 100% accurate ground truth data and other improvements in input data as suggested in the discussion, the CV method could produce classifications with over 80% accuracy. A few of the datasets already surpassed 80% overall accuracy using the CDL ground truth, while at the same time examination of the main sources of error shows how improved image collection dates and ground truth data could easily lead to improved classifications.

The upcoming NISAR mission combined with field campaigns to collect more accurate ground truth data provides an opportunity to solve both of these issues. It would also allow for further investigations of climatic influences on CV values. Using the CV method with NISAR data makes possible the use of L-band data for consistent global agricultural monitoring.

## Acknowledgements

This work was supported by NASA Headquarters under a grant for the NASA Science Definition Team for the NASA ISRO Synthetic Aperture Radar (NISAR) Mission (NNX16AK59G). The authors would also like to thank the Japan Aerospace Exploration Agency's (JAXA) Kyoto and Carbon Initiative and the Alaska Satellite Facility for providing the ALOS data.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.jag.2017.12.014.

## References

Alaska Satellite Facility Engineering Group, 2015. ASF Radiometrically Terrain Corrected ALOS PALSAR Products: Product Guide.

Bartholomé, E., Belward, A.S., 2005. GLC2000: a new approach to global land cover mapping from Earth observation data. Int. J. Remote Sens. 26, 1959–1977. http://dx.doi.org/10.1080/01431160412331291297.

Becker-Reshef, I., Justice, C.O., Sullivan, M., Vermote, E., Tucker, C., Anyamba, A., Small, J., Pak, E., Masuoka, E., Schmaltz, J., Hansen, M., Pittman, K., Birkett, C., Williams, D., Reynolds, C., Doorn, B., 2010. Monitoring global croplands with coarse resolution earth observations: the Global Agriculture Monitoring (GLAM) project. Remote Sens. 2, 1589–1609. http://dx.doi.org/10.3390/rs2061589.

Biradar, C.M., Thenkabail, P., Noojipady, P., Li, Y., Dheeravath, V., Turral, H., Velpuri, M., Gumma, M.K., Gangalakunta, O.R.P., Cai, X.L., Xiao, X., Schull, M.A., Alankara, R., Gunasinghe, S., Mohideen, S., 2009. A global map of rainfed cropland areas (GMRCA) at the end of last millennium using remote sensing. Int. J. Appl. Earth Obs. Geoinf. 11, 114–129. http://dx.doi.org/10.1016/j.jag.2008.11.002.

Blaes, X., Vanhalle, L., Defourny, P., 2005. Efficiency of crop identification based on optical and SAR image time series. Remote Sens. Environ. 96, 352–365. http://dx.doi.org/10.1016/j.rse.2005.03.010.

Bontemps, S., Defourny, P., Van Bogaert, E., Arino, O., Kalogirou, V., Ramos Perez, J., 2011. Globcover 2009 Products Description and Validation Report. European Space Agency (ESA)/Universite catholique de Louvain (UCL).

Boryan, C.G., Yang, Z., Mueller, R., Craig, M., 2011. Monitoring US agriculture: the US department of agriculture, national agricultural statistics service, cropland data layer program. Geocarto Int. 26, 341–358. http://dx.doi.org/10.1080/10106049.2011.562309.

Bouvet, A., Le Toan, T., 2011. Use of ENVISAT/ASAR wide-swath data for timely rice fields mapping in the Mekong River Delta. Remote Sens. Environ. 115, 1090–1101. http://dx.doi.org/10.1016/j.rse.2010.12.014.

Chakraborty, M., Patnaik, C., Panigrahy, S., Parihar, J.S., 2006. Kuligowski, R.J., Parihar, Jai S., Saito, G. (Eds.), Monitoring of Wet Season Rice Crop at State and National Level in India Using Multidate Synthetic Aperture Radar Data 641103. http://dx.doi.org/10.1117/12.693900.

Champagne, C., McNairn, H., Daneshfar, B., Shang, J., 2014. A bootstrap method for assessing classification accuracy and confidence for agricultural land use mapping in Canada. Int. J. Appl. Earth Obs. Geoinf. 29, 44–52. http://dx.doi.org/10.1016/j.jag.2013.12.016.

Cihlar, J., Pultz, T.J., Gray, A.L., 1992. Change detection with synthetic aperture radar. Int. J. Remote Sens. 13, 401–414.

Daly, C., Widrlechner, M.P., Halbleib, M.D., Smith, J.I., Gibson, W.P., 2012. Development of a new USDA plant hardiness zone map for the United States. J. Appl. Meteorol. Climatol. 51, 242–264. http://dx.doi.org/10.1175/2010JAMC2536.1.

Deschamps, B., McNairn, H., Shang, J., Jiao, X., 2012. Towards operational radar-only crop type classification: comparison of a traditional decision tree with a random forest classifier. Can. J. Remote Sens. 38, 60–68. http://dx.doi.org/10.5589/m12-012.

Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recognit. Lett. 27, 861–874. http://dx.doi.org/10.1016/j.patrec.2005.10.010.

Ferrazzoli, P., Paloscia, S., Pampaloni, P., Schiavon, G., Sigismondi, S., Solimini, D., 1997. The potential of multifrequency polarimetric SAR in assessing agricultural and arboreous biomass. IEEE Trans. Geosci. Remote Sens. 35, 5–17. http://dx.doi.org/10.1109/36.551929.

Fisette, T., Davidson, A., Daneshfar, B., Rollin, P., Aly, Z., Campbell, L., 2014. In: Annual Space-based Crop Inventory for Canada: 2009–2014. Presented at the Geoscience and Remote Sensing Symposium (IGARSS). 2014 IEEE International, IEEE, Quebec City, Canada. http://dx.doi.org/10.1109/IGARSS.2014.6947643.

Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X., 2010. MODIS Collection 5 global land cover: algorithm refinements and characterization of new datasets. Remote Sens. Environ. 114, 168–182. http://dx.doi.org/10.1016/j.rse.2009.08.016.

Gong, P., Wang, J., Zhao, Yongchao, Zhao, Yuanyuan, Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S., Li, C., Li, X., Fu, W., Liu, C., Xu, Y., Wang, X., Cheng, Q., Hu, L., Yao, W., Zhang, Han, Zhu, P., Zhao, Z., Zhang, Haiyang, Zheng, Y., Ji, L., Zhang, Y., Chen, H., Yan, A., Guo, J., Yu, L., Wang, L., Liu, X., Shi, T., Zhu, M., Chen, Y., Yang, G., Tang, P., Xu, B., Giri, C., Clinton, N., Zhu, Z., Chen, Jin, Chen, Jun, 2013. Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data. Int. J. Remote Sens. 34, 2607–2654. http://dx.doi.org/10.1080/01431161.2012.748992.

Haldar, D., Das, A., Mohan, S., Pal, O., Hooda, R.S., Chakraborty, M., 2012. Assessment of L-band SAR data at different polarization combinations for crop and other landuse classification. Progress Electromagn. Res. B 36, 303–321.

Hoekman, D.H., Vissers, M.A.M., 2003. A new polarimetric classification approach evaluated for agricultural crops. IEEE Trans. Geosci. Remote Sens. 41, 2881–2889. http://dx.doi.org/10.1109/TGRS.2003.817795.

Hong, G., Zhang, A., Zhou, F., Brisco, B.G., 2014. Integration of optical and synthetic aperture radar (SAR) images to differentiate grassland and alfalfa in Prairie area. Int. J. Appl. Earth Obs. Geoinf. 28, 12–19. http://dx.doi.org/10.1016/j.jag.2013.10.003.

Jayne, T.S., Rashid, S., 2010. The value of accurate crop production forecasts. In: International Development Working Papers, #108. Presented at the Fourth African Agricultural Markets Program (AAMP) Policy Symposium. Lilongwe, Malawi.

Jiao, X., Kovacs, J.M., Shang, J., McNairn, H., Walters, D., Ma, B., Geng, X., 2014. Object-oriented crop mapping and monitoring using multi-temporal polarimetric RADARSAT-2 data. ISPRS J. Photogramm. Remote Sens. 96, 38–46. http://dx.doi.org/10.1016/j.isprsjprs.2014.06.014.

Kendall, M.G., Stuart, A., 1969. The Advanced Theory of Statistics, 3rd ed. Charles Griffin & Company Limited, London.

Lucas, R.M., Armston, J., Fairfax, R., Fensham, R., Accad, A., Carreiras, J.M.B., Kelley, J., Bunting, P., Clewley, D., Bray, S., Metcalfe, D., Dwyer, J., Bowen, M., Eyre, T., Laidlaw, M., Shimada, M., 2010. An evalulation of the ALOS PALSAR L-band backscatter – above ground biomass relationship queensland, Australia: impacts of surface moisture condition and vegetation structure. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 3, 576–593. http://dx.doi.org/10.1109/JSTARS.2010.2086436.

Ma, X., Shen, H., Yang, J., Zhang, L., Li, P., 2014. Polarimetric-spatial classification of SAR images based on the fusion of multiple classifiers. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7, 961–971. http://dx.doi.org/10.1109/JSTARS.2013.2265331.

Matton, N., Sepulcre Cant, G., Waldner, F., Valero, S., Morin, D., Inglada, J., Arias, M., Bontemps, S., Koetz, B., Defourny, P., 2015. An automated method for annual cropland mapping along the season for various globally-distributed agrosystems using high spatial and temporal resolution time series. Remote Sens. 7, 13208–13232. http://dx.doi.org/10.3390/rs71013208.

McNairn, H., Shang, J., 2016. A review of multitemporal synthetic aperture radar (SAR) for crop monitoring. In: Ban, Y. (Ed.), Multitemporal Remote Sensing: Methods and Applications, Remote Sensing and Digital Image Processing. Springer International Publishing, Cham, pp. 317–340. http://dx.doi.org/10.1007/978-3-319-47037-5_15.

McNairn, H., Shang, J., Jiao, X., Champagne, C., 2009. The contribution of ALOS PALSAR multipolarization and polarimetric data to crop classification. IEEE Trans. Geosci. Remote Sens. 47, 3981–3992. http://dx.doi.org/10.1109/TGRS.2009.2026052.

NASA, 2012. Data Rights & Related Issues.

National Oceanic and Atmospheric Administration, National Weather Service (NOAA/NWS), 2017. NOWData – NOAA Online Weather Data [Governmental Page]. Boise, ID, Pocatello, ID, Dodge City, KS, Bismarck, ND, Pendleton, OR, Memphis, TN, and Corpus Christi, TX National Weather Service Forecast Offices. Accessed from http://w2.weather.gov/climate/index.php?wfo=boi (Boise), http://w2.weather.gov/climate/index.php?wfo=pih (Pocatello), http://w2.weather.gov/climate/index.php?wfo=ddc (Dodge City), http://w2.weather.gov/climate/index.php?wfo=bis (Bismarck), http://w2.weather.gov/climate/index.php?wfo=pdt (Pendleton), http://w2.weather.gov/climate/index.php?wfo=meg (Memphis), and http://w2.weather.gov/climate/index.php?wfo=crp (Corpus Christi).

Nelson, A., Setiyono, T., Rala, A.B., Quicho, E.D., Raviz, J.V., Abonete, P.J., Maunahan, A.A., Garcia, C.A., Bhatti, H.Z.M., Villano, L.S., Thongbai, P., Holecz, F., Barbieri, M., Collivignarelli, F., Gatti, L., Quilang, E.J.P., Mabalay, M.R.O., Mabalot, P.E., Barroga, M.I., Bacong, A.P., Detoito, N.T., Berja, G.B., Varquez, F., Wahyunto, Kuntjoro D., Murdiyati, S.R., Pazhanivelan, S., Kannan, P., Mary, P.C.N., Subramanian, E., Rakwatin, P., Intrman, A., Setapayak, T., Lertna, S., Minh, V.Q., Tuan, V.Q., Duong, T.H., Quyen, N.H., Kham, D.V., Hin, S., Veasna, T., Yadav, M., Chin, C., Ninh, N.H., 2014. Towards and operation SAR-based rice monitoring system in Asia: examples from 13 demonstration sites across Asia in the RIICE project. Remote Sens. 6, 10773–10812. http://dx.doi.org/10.3390/rs61110773.

Peel, M.C., Finlayson, B.L., McMahon, T.A., 2007. Updated world map of the Koppen-

Geiger climate classification. Hydrol. Earth Syst. Sci. 11, 1633–1644.
Pittman, K., Hansen, M.C., Becker-Reshef, I., Potapov, P.V., Justice, C.O., 2010. Estimating global cropland extent with multi-year MODIS data. Remote Sens. 2, 1844–1863. http://dx.doi.org/10.3390/rs2071844.
Ramankutty, N., Evan, A.T., Monfreda, C., Foley, J.A., 2008. Farming the planet : 1. Geographic distribution of global agricultural lands in the year 2000. Glob. Biogeochem. Cycles 22. http://dx.doi.org/10.1029/2007GB002952.
Rosen, P.A., Hensley, S., Shaffer, S., Veilleux, L., Chakraborty, M., Misra, T., Bhan, R., Raju Sagi, V., Satish, R., 2015. The NASA-ISRO SAR mission ? an international space partnership for science and societal benefit. In: Radar Conference (RadarCon). 2015 IEEE. Presented at the 2015 IEEE Radar Conference (RadarCon), IEEE, Arlington, VA. pp. 1610–1613. http://dx.doi.org/10.1109/RADAR.2015.7131255.
Rosenqvist, A., Shimada, M., Suzuki, S., Ohgushi, F., Tadono, T., Watanabe, M., Tsuzuku, K., Watanabe, T., Kamijo, S., Aoki, E., 2014. Operational perfomance of the ALOS global systematic acquisition strategy and observation plans for ALOS-2 PALSAR-2. Remote Sens. Environ. 155, 3–12. http://dx.doi.org/10.1016/j.rse.2014.04.011.
Sanchez, N., 2014. NISAR Project Science Requirements – Initial Release D-76290.
Scherr, S.J., Sthapit, S., 2009. Mitigating Climate Change Through Food and Land Use (No. 179), Worldwatch Reports. Ecoagriculture Partners and Worldwatch Institute, Washington D.C.
Siqueira, P., Saatchi, S.S., 2015. Agriculture & Ecosystems InSAR Inundation and Crop Area.
Siqueira, P., 2016. Interoffice Memorandum – Use of ESA's AGRISAR and Sentinel-1A Data for Validating NISAR's Crop Area Requirement.
Skriver, H., Mattia, F., Satalino, G., Balenzano, A., Pauwels, V.R.N., Verhoest, N.E.C., Davidson, M., 2011. Crop classification using short-Revisit multitemporal SAR data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 4, 423–431. http://dx.doi.org/10.1109/JSTARS.2011.2106198.
Small, D., 2011. Flattening gamma: radiometric terrain correction for SAR imagery. IEEE Trans. Geosci. Remote Sens. 49, 3081–3093. http://dx.doi.org/10.1109/TGRS.2011.2120616.
Stefanski, J., Chaskovskyy, O., Waske, B., 2014. Mapping and monitoring of land use changes in post-soviet western Ukraine using remote sensing data. Appl. Geogr. 55, 155–164. http://dx.doi.org/10.1016/j.apgeog.2014.08.003.
Tso, B., Mather, P.M., 1999. Crop discrimination using multi-temporal SAR imagery. Int. J. Remote Sens. 20, 2443–2460. http://dx.doi.org/10.1080/014311699212119.
Waldner, F., Fritz, S., Di Gregorio, A., Defourny, P., 2015. Mapping priorities to focus cropland mapping activities: fitness assessment of existing global, regional and national cropland maps. Remote Sens. 7, 7959–7986. http://dx.doi.org/10.3390/rs70607959.
Whitcraft, A.K., Becker-Reshef, I., Killough, B.D., Justice, C.O., 2015. Meeting earth observation requirements for global agricultural monitoring: an evaluation of the revisit capabilities of current and planned moderate resolution optical earth observing missions. Remote Sens. 7, 1482–1503. http://dx.doi.org/10.3390/rs70201482.
Yu, L., Wang, J., Clinton, N., Xin, Q., Zhong, L., Chen, Y., Gong, P., 2013. FROM-GC: 30 m global cropland extent derived through multisource data integration. Int. J. Digital Earth 6, 521–533.
Yusoff, N.M., Muharam, F.M., Takeuchi, W., Darmawan, S., Razak, M.H.A., 2017. Phenology and classification of abandoned agricultural land based on ALOS-1 and 2 PALSAR multi-temporal measurements. Int. J. Digit. Earth 10, 155–174. http://dx.doi.org/10.1080/17538947.2016.1216615.