ANNUAL SPACE-BASED CROP INVENTORY FOR CANADA: 2009-2014

Fisette, T., Davidson, A., Daneshfar, B., Qiao, C., Rollin, P., Aly, Z., Campbell, L.

Agriculture and Agri-Food Canada, Ottawa, Canada

ABSTRACT

Starting in 2009, Agriculture and Agri-Food Canada (AAFC) began the process of generating annual crop type digital maps using satellite imagery. Focusing on the Prairie Provinces in 2009 and 2010, a Decision Tree (DT) based methodology was applied using optical (Landsat, Resourcesat-1, DMC, SPOT) and radar (Radarsat-2) imagery. Starting with the 2011 growing season, this activity was extended to other provinces in support of a national crop inventory. Since then, AAFC has consistently delivered an annual crop inventory that is close to the overall target accuracy of 85% at the national scale. However, from one province to another, and from one year to another, the classification accuracy is not uniform. It varies depending on satellite data availability and training site distribution. In an effort to improve mapping quality, the methodology is constantly evolving and new classifiers are under investigation.

***Index Terms*—** Land Cover, Land Use, Remote Sensing, Agriculture, Classification

1. INTRODUCTION

Understanding the state and trends in agriculture production is essential to combat both short-term and long-term threats to stable and reliable access to food for all, and to ensure a profitable agricultural sector. However, because Canada's agricultural landscape is extensive and diverse, our ability to manage it is only as good as the information available to make informed decisions. Space-based Earth Observation (EO) can deliver cost-effective, timely and accurate information to better support policies, programs, performance measurement and market access.

AAFC is in a unique position to capitalize on the integration of EO technology, ground observation data and other monitoring systems to provide information relating to agricultural production. AAFC is a leader in the development and use of EO technologies for this type of agricultural assessment, and has already demonstrated that these technologies can deliver cost-effective, timely and accurate information [1,2].

The AAFC's Annual Space-Based Crop Inventory for Canada (Fig.1) provides high quality information on the location, extent and changes of Canadian crops. The inventory, first brought online in 2009, has had an impact on the Canadian agriculture sector and beyond. Within AAFC, the crop inventory is an important foundational data source for a number of activities, and it will support the next generation of environmental indicators, and improve business risk program implementation and verification (i.e. Agri-Recovery decision making, drought impacts, and excess wetness).

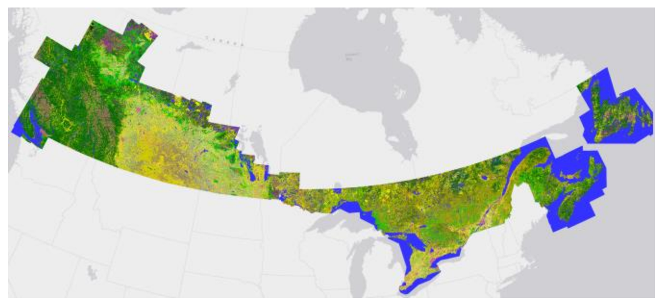

Fig. 1. The 2013 AAFC crop inventory.

Externally, the crop inventory has had a major impact on competitiveness by supporting Canadian canola producers' access to the European bio-fuel feedstock markets, with an estimated potential of $500M/yr. The Province of Alberta uses the inventory in its land use framework, to monitor the impact of land fragmentation on agricultural producers. Most recently the inventory is being regarded as a cost effective source of land use data to replace survey-based information that will no longer be supported by Statistics Canada.

2. METHODOLOGY

Successful crop identification relies on image acquisitions from multiple sensors during key crop phenological stages. Multi-temporal optical data are the primary data source for crop classification because the NIR/SWIR channels are highly sensitive to vegetation changes. Over a growing season, at least three optical images are required to successfully identify crops. Non-ideal weather conditions during key growth stages can lead to gaps in the image record. To fill these gaps dual-polarization RADARSAT-2 (RSAT2) data are used throughout the growing season. The radar data also brings additional information as it is much more sensitive to plant structure than optical data. The

ScanSAR mode, with its large swath (300 km) and moderate resolution (50 m), fits the agricultural landscape of the Prairie Provinces. Elsewhere, the finer resolution of the Wide mode (25 m) is better suited to narrow fields. The number of images used varies from year to year depending on available sensors (Fig.2). DMC (Disaster Monitoring Constellation) optical images have the advantage of covering very wide swath (650 km) at a 32/22 m resolution but its price remains an obstacle to its operational use. In 2012, due to the lack of affordable data, AAFC had to rely mostly on RADARSAT-2 data. To maintain crop map accuracy levels comparable to 2011, the number of radar images ordered more than doubled from 570 in 2011 to 1380 in 2012. By 2013, this number decreased slightly with the arrival of Landsat-8 data early in the growing season.

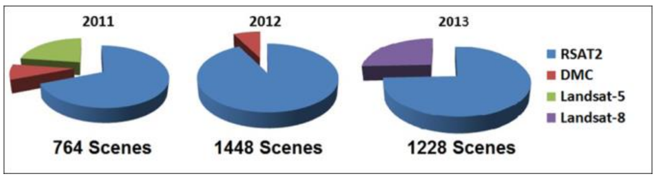

Fig. 2. Number and types of images used for classification (2011 to 2013).

In 2012, AAFC processed more than 2500 combinations of imagery in our classifier. Results have shown that a classification containing only one or two optical images can have its overall accuracy improved by up to 16% by adding dual-polarization radar images [3].

Ground truth information used in the model training and validation is critical to successful crop classification. Currently, crop insurance data are the most accurate, detailed and complete sources of information for geospatial crop type information in Canada. As such, AAFC gets data from crop insurance agencies in Alberta, Saskatchewan and Quebec. For provinces where insurance data cannot be accessed, ground-truth information is provided by point observations from AAFC staff or other provincial sources.

The mapping process doesn't require to mosaic multi-date optical data so there is no need for atmospheric correction. If not already geometrically corrected, images were orthorectified using a 3-D multi-sensor physical model developed at the Canada Centre for Remote Sensing [4] and implemented in PCI Geomatica software. An automated cloud and shadow masking technique [5] is applied to every optical image. A gamma maximum-a-posteriori filter is applied on radar data to remove frequency noise (speckle) and resampled to 16-bit to reduce processing time. Along-track images of the same date are mosaicked together. This process allows the creation of large classification regions with more training sites.

The country is divided into regions that are classified independently. Each of these regions combines several optical and radar dates. Around 130 regions are defined nationally and the combination of imagery per region depends on many factors including: image overlap, cloud cover, agricultural extent and training site distribution. Classifications are performed on a region-by-region basis because the dynamic nature of crop rotations, crop growth and harvest patterns create significant reflectance differences between adjacent satellite scenes within the temporal period encompassed by scene availability. Time series images are used for each classification in order to separate the various crops classes, as they have varying spectral characteristics over the growing season.

The DT method, as implemented in See5 software, is a multivariate model based on a set of decision rules defined by combinations of features and a set of linear discriminant functions that are applied at each test node. Decision boundaries and coefficients for the linear discriminate function are estimated empirically from the training data. The DT method was chosen because of its ability to handle discrete data, its processing speed, its independence of the distribution of class signatures, its interpretable classification rules [6,7] and its cost effectiveness. Advanced options such as pruning and boosting have also been incorporated into the decision tree classification process to improve the accuracy of the algorithm.

The classification process is divided into three iterations. The first iteration is updating the AAFC circa 2000 land cover to the current year. A random selection of training sites are gathered from across the land cover map and used as inputs to the classifier. The second iteration uses the current year training sites as inputs to the classifier, and is mapping only two classes: agriculture and grassland. The final iteration is the actual crop classification that is restricted to the agriculture extent of the previously created map. Within a region, clouds, overlapping and non-overlapping areas must be processed independently as they will generate different rule sets. One of the See5 outputs is a thematic layer partitioning the scene into homogenous zones representing these areas. These zones contain crop accuracy values estimated from the validation data.

For each classified region, the multi-temporal optical bands are inputted in to the eCognition segmentation algorithm that derives polygons (object primitives). A spatial scale of 10 was found to be a good compromise between the number of objects and individual field representation. The polygons are imported into the ESRI ArcGIS Spatial Analyst module and a majority zonal statistic was calculated on the per pixel classification. This assigned the majority class value within each object primitive to the entire polygon. That step improves the overall accuracy by about 5%. Classification regions are then mosaicked together through an automated process that prioritizes zones with higher accuracies in overlapped areas.

3. DISCUSSION

For AAFC, assessing map accuracy over the years and across Canada is important to improve the quality of the product. For such a large scale however, it is not an easy task to describe the accuracy. The error matrix is the most common way of expressing the accuracy of image classifications and has been criticized for not providing any indication of the spatial distribution of errors [8]. The accuracy evaluation of the annual crop inventory is done through several approaches.

3.1. Product Accuracy

The overall crop accuracy shows a lot of variation from one province to the other (Table 1). This can be explained by several factors that are encountered during the acquisition and processing of the data. Firstly, the ground data acquisition is made on a per-province basis. The density and quality of training sites is not homogeneous across Canada. Accuracies for Provinces where we have access to crop insurance data are usually very stable through time. This is the case for Alberta, Saskatchewan and Quebec, with yearly accuracy variation not exceeding 6%. Secondly, in some areas, cloud cover is significant enough to limit the amount of available spectral information. Additionally, crop classes such as the cereal sub-categories (barley, oats, wheat, etc.) have been aggregated into a single cereal class in all provinces except for Alberta, Saskatchewan and Quebec. This results in class discontinuities between provinces. At the national scale, when combining all provinces together, the year 2012 has the lowest crop overall accuracy (83.9%). The poor optical data availability for that year, did not allowed us to reach our 85% accuracy goal.

Table 1. Overall crop accuracies per province and per year.

Provinces	Overall Accuracy (Per Year)				
	2009	2010	2011	2012	2013
PEI	-	-	67.5%	78.7%	87.9%
NB	-	-	88.1%	88.0%	87.3%
NS	-	-	64.2%	89.9%	74.2%
QC	-	-	81.4%	81.8%	87.5%
ON	-	-	80.8%	76.2%	88.2%
MB	80.0%	-	79.0%	85.9%	85.4%
SK		88.3%	87.1%	82.4%	86.5%
AB			87.7%	88.4%	89.9%
BC	-	-	-	-	79.2%
Canada	80.0%	88.3%	85.3%	83.9%	86.0%

During the classification process, the overall crop accuracy value is calculated for each unique imagery date combination. These accuracies are fundamental to the mosaic process as they prioritize higher accuracy regions over lower accuracy regions. When mapped, that accuracy data provides a very detailed representation of the spatial variations in crop accuracy. Figure 3 shows this information at the national scale for the year 2013. The accuracy is very heterogeneous across the country, with the exception of a large area over the provinces of Alberta and Saskatchewan. For this area, the availability of cloud-free landsat-8 imagery and the proper distribution of training sites, contributes to the accuracy's consistency which varies between 85% and 90%. Elsewhere, in most cases, regions with low accuracy can be explained by the poor availability of Landsat-8 data (presence of clouds). This is the case for the Maritime Provinces, where some crop classes had to be combined into generic classes in order to improve the accuracy. The analysis of these maps over several years will allow AAFC to identify recurring patterns and thus adjust our mapping approach.

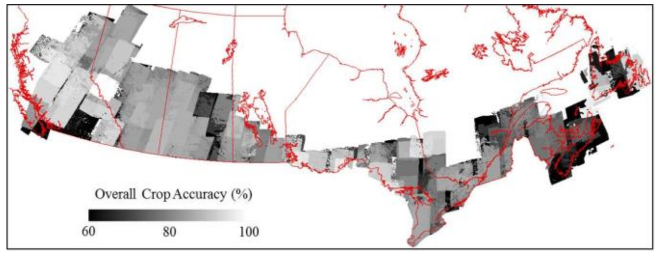

Fig. 3. Spatial distribution of the overall crop accuracy (year 2013).

3.2. Comparison with Census of Agriculture

Census of Agriculture data collected by Statistics Canada in 2011 was compared to our crop inventory for the same year. For majors crops (forage, cereals, oilseeds), the cultivated areas from the AAFC inventory are 7% to 14% higher than the census (Fig.4). For less common crops, with the exception of corn and fruits, the area estimated by AAFC can be greatly underestimated: 71% for fallow and 63% for pulses (Fig.5). At the national scale, the AAFC inventory overestimates the agriculture area by 15% compared to the 2011 Census of Agriculture.

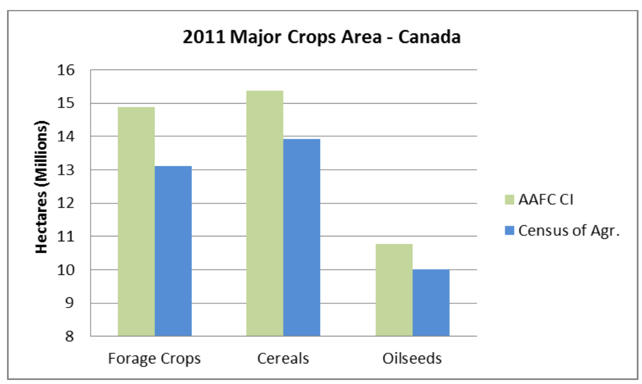

Fig. 4. Areas comparison between the AAFC Crop Inventory (CI) and the Statistics Canada Census of Agriculture for the major crops in 2011.

Those differences might be explained by the number and distribution of training sites. Quality of the results depends crucially on the adequacy of the training data to represent the classes [9]. The impact of the distribution of training sites on the accuracy of a classification has been the subject of a joint study between Statistics Canada and AAFC. The results show that the distribution of training sites greatly affected the classification results. In general, the more a class is represented in the training sites, the more it will be over-represented in the classification results. This will occur to the detriment of marginal crops that are underrepresented. These results will help AAFC to develop a recalibration approach to optimize the training sites distribution.

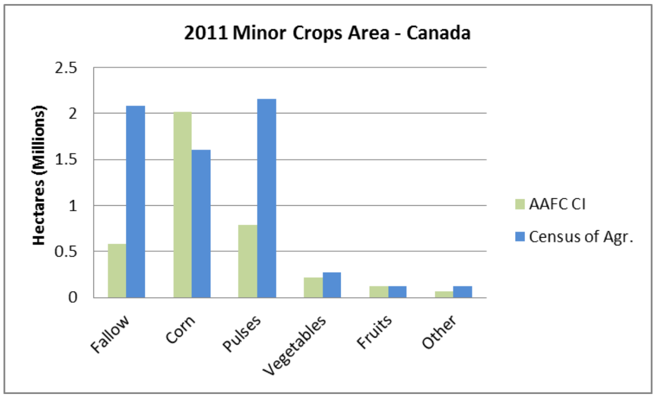

Fig. 5. Areas comparison between the AAFC Crop Inventory (CI) and the Statistics Canada Census of Agriculture for the minor crops in 2011.

3.3. Future Implementation

Despite the demonstrated success of the Crop Inventory, there are areas of development that must be addressed if it is going to be able to meet the future needs of AAFC and its clients in a cost-effective and computationally efficient manner. Recently, new classifier algorithms, such as the Random Forest (RF) classifier, have become available. The RF outperformed the DT by classifying 18 times faster [10]. For six regions across the country, the RF classifier was tested against the AAFC DT classifier. In all cases, the RF algorithm was more accurate than the DT.

There is a growing demand for the crop inventory to be made available immediately following the growing season. Many users would also like to see an estimated inventory published within the season. Achieving these objectives will require the optimization of our classification process that currently takes up to a week for a single region. A new and fully automated crop classifier that should significantly reduce production time is under production. Ground truth data collection strategy will need to be reviewed as training sites in some provinces are made available late fall following the growing season. Over the next 4 years, modifications to our classification system will be made to ingest and process new satellite data such as the RADARSAT Constellation Mission (RCM) and the European Space Agency new family of satellites called Sentinels.

4. CONCLUSION

A Decision Tree (DT) method has been developed to successfully classify Canadian agricultural lands. On average, since 2011, AAFC crop maps are 85% accurate. Our intent is not to build a static methodology but rather to constantly improve it as our classification work evolves. Comparison with other datasets such as the Census of Agriculture contribute to a better understanding of the strengths and weaknesses of the crop map products, and helps target AAFC effort to enhance it. Integration of other earth observation imagery and geospatial data within the classifier will continue to be explored. As part of the Canadian federal government commitment to open data, the entire datasets is uploaded to http://data.gc.ca.

11. REFERENCES

[1] H. McNairn, A. Merzouki, A. Pacheco and J. Fitzmaurice, “Monitoring Soil Moisture to Support Risk Reduction for the Agriculture Sector Using RADARSAT-2,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(3), pp. 252-259, 2012.

[2] H. McNairn, C. Champagne, J. Shang, D. Holmstrom and G. Reichert, “Integration of optical and Synthetic Aperture Radar (SAR) imagery for delivering operational annual crop inventories,” *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(5), pp. 434-449, 2009.

[3] T. Fisette, P. Rollin, Z. Aly, L. Campbell, B. Daneshfar, P. Filyer, A. Smith, A. Davidson, J. Shang and I. Jarvis, “AAFC annual crop inventory,” *Second International Conference on Agro-Geoinformatics, Fairfax, VA*, pp. 270-274, August 2013.

[4] T. Toutin, “Multi-source data fusion with an integrated and unified geometric modelling,” *EARSeL advances in remote sensing*, 4(2), pp. 118-129, 1995.

[5] Z. Zhu and C. E. Woodcock, “Object-based cloud and cloud shadow detection in Landsat imagery,” *Remote Sensing of Environment*, 118(0), pp. 83-94, 2012.

[6] M.A. Friedl and C.E. Brodley, “Decision tree classification of land cover from remotely sensed data,” *Remote Sensing of Environment*, 61(3), pp. 399-409, 1997.

[7] M. Pal and P.M. Mather, “An assessment of the effectiveness of decision tree methods for land cover classification,” *Remote Sensing of Environment*, 86(4), pp. 554-565, 2003.

[8] A. Comber, P. Fisher, C. Brunsdon and A. Khmag, “Spatial Analysis or remote sensing image classification accuracy,” *Remote Sensing of Environment*, 127, pp. 237-246, 2012.

[9] M. Pal and P. M. Mather, “An assessment of the effectiveness of decision tree methods for land cover classification,” *Remote Sensing of Environment*, 86(4), pp. 554-565, 2003.

[10] B. Deschamps, H. McNairn, J. Shang and X. Jiao, “Towards operational radar-only crop type classification: comparison of a traditional decision tree with a random forest classifier,” *Canadian Journal of Remote Sensing*, 38(1), pp. 60-68, 2012.