**Report: Comparison Between MASC and ACI Agricultural Data for Manitoba (2017–2024)**

**Overview**  
  
We learned that the MASC (Manitoba Agricultural Services Corporation) dataset and the AAFC Annual Crop Inventory (ACI) both provide large-scale accounts of crop acreage and type across Manitoba, but they differ substantially in definitions, data sources, and classification methodologies. By comparing imputed MASC totals with ACI summaries for 2017–2024, we identified consistent and measurable discrepancies in both total agricultural area and crop-specific representation.  
  
MASC provides ground-truth administrative data collected through farm insurance reporting and direct surveys. ACI provides remote-sensing classifications derived from optical and radar imagery combined with supervised learning algorithms. The two are not designed to measure the same quantity, but their totals can still be compared to understand bias and classification performance.

**Total Agricultural Area**  
  
We learned that ACI consistently reports higher total agricultural acreage than MASC.  
  
For 2024:  
Total MASC acres: 9,455,541  
Total ACI acres: 13,070,690  
Difference (ACI − MASC): 3,615,148 acres  
  
This equates to a roughly 38 percent overestimation of agricultural land by ACI relative to MASC.  
  
When compared to Statistics Canada’s Census of Agriculture, which reported 17.1 million acres of total farmland in Manitoba in 2021, MASC’s reported acreage corresponds well with the share used for crop production (approximately 68 percent of total farmland, or 11.6 million acres). This supports the interpretation that MASC measures actively cultivated cropland, while ACI includes additional managed land classes such as pasture, hay, and fallow.

**Literature Context**  
  
Fisette et al. (2013) reported that “At the national scale, the AAFC inventory overestimates the agriculture area by about 15% compared to the 2011 Census of Agriculture.” This early study of ACI methodology showed a consistent upward bias in the classified agricultural extent, even at the national level. Later AAFC documentation continues to note that “training data are derived primarily from crop insurance datasets and ground reference surveys... the number and geographic distribution of reference samples reflect the spatial and thematic coverage of these sources” (AAFC, 2019).  
  
This means that major crops, such as canola and spring wheat, are sampled in proportion to their prevalence in provincial crop insurance data, which leads to a non-stratified sampling scheme that favors dominant classes. The resulting decision tree classifier performs very well on high-prevalence classes but may overgeneralize in marginal or spectrally similar regions.

**Crop-Level Discrepancies**  
  
We learned that the overestimation pattern varies across crop classes.  
  
In 2024, the top-level comparison showed:  
Canola/rapeseed: +1,035,406 acres difference (ACI overestimation of 26.7 percent)  
Spring wheat: +1,229,600 acres difference (ACI overestimation of 36.4 percent)  
Soybeans: +401,337 acres difference (ACI overestimation of 29.7 percent)  
Pasture/forages: +694,434 acres difference (ACI overestimation of 94.0 percent)  
Corn: +186,866 acres difference (ACI overestimation of 30.3 percent)  
  
Conversely, underrepresented or misclassified classes included:  
Mustard: −3,776 acres (−69.6 percent)  
Flaxseed: −9,389 acres (−23.0 percent)  
Millet: −2,910 acres (−49.3 percent)  
Triticale: −2,831 acres (−77.6 percent)  
  
These findings suggest that dominant crops show high absolute overrepresentation, while minority crops show higher relative percentage errors. The absolute difference is driven by total land area, while the percentage difference reflects the classifier’s relative performance on smaller or underrepresented classes.

**Relationship Between Absolute and Relative Error**  
  
We learned that overestimation as a percentage does not correlate with crop dominance, while absolute overestimation does.  
  
Major crops such as canola, wheat, and soybeans do not exhibit the highest percentage error but dominate the total discrepancy in absolute acres. Smaller or less common crops, such as mustard or millet, show very high percentage differences, but their total acreage is too small to meaningfully affect the total provincial discrepancy.  
  
This supports the conclusion that ACI’s training process achieves strong performance on well-represented classes but struggles to generalize on minority classes with few labeled samples. The behavior is consistent with non-stratified sampling, where training data for each class are drawn in proportion to their occurrence, not equally across classes.

**Algorithmic and Structural Causes**  
  
We learned that ACI’s bias toward overestimating major crops likely originates from two intertwined factors:  
  
1. Structural bias in definition:  
 ACI defines “agricultural land” broadly, including perennial vegetation, pasture, and managed hay. MASC defines it narrowly as cropland under production. This definitional difference inflates ACI totals even if classification were perfect.  
  
2. Algorithmic bias from unbalanced training data:  
 ACI’s classification model (a decision tree algorithm, as documented by AAFC) uses a larger number of training samples from dominant crop classes. According to Fisette et al. (2013) and Davidson et al. (2017), “sampling reflected the proportional availability of crop insurance and field survey data, which were dominated by the major crops.”  
 This imbalance leads the model to favor majority classes when faced with ambiguous spectral signatures. When a classifier is uncertain, it tends to assign pixels to the class it has seen most often in training, resulting in systematic overclassification of dominant crops such as canola or spring wheat.

**Connection to Previous Findings**  
  
The 2011 overestimation of agricultural area by 15 percent (Fisette et al., 2013) and the 2024 overestimation of roughly 27–30 percent in Manitoba suggest that the general magnitude of the discrepancy has persisted over time despite algorithmic and sensor improvements. The persistence of this bias implies that it is primarily due to definitional and data-source differences rather than technical error.

**Magnitudes by Year**

**A screenshot of a graph

AI-generated content may be incorrect.**

**A screenshot of a graph

AI-generated content may be incorrect.**

A green and red graph

AI-generated content may be incorrect.

**Conclusion**  
  
We learned that discrepancies between ACI and MASC data in Manitoba arise from both definitional and algorithmic factors. ACI consistently reports more agricultural land because its classification includes a broader set of vegetated surfaces and because its decision tree model was trained with a disproportionate number of samples from dominant crop classes.  
  
While percentage differences are higher for minority crops, the large absolute differences for dominant crops such as canola and spring wheat have the greatest effect on total acreage discrepancies. The same structural overestimation has been documented in national-scale studies since ACI’s inception, suggesting that it is a persistent characteristic of the dataset rather than a regional or temporal anomaly.  
  
The Manitoba case reinforces the importance of cross-validating remote-sensing classifications against ground-based administrative data such as MASC to contextualize systematic biases and better interpret provincial-scale agricultural statistics.

**References**

Fisette, T. (2012). Canadian Space-Based Crop Inventory (CSRS 2012 presentation). Agriculture and Agri-Food Canada, Earth Observation Team.

Fisette, T. (2019, August). Annual Crop Inventory: Agriculture and Agri-Food Canada (Guest lecture slides, Version 3). Agriculture and Agri-Food Canada.

Fisette, T., Davidson, A., Daneshfar, B., Qiao, C., Rollin, P., Aly, Z., & Campbell, L. (2014). Annual space-based crop inventory for Canada: 2009–2014. IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2014), 3089–3096. Agriculture and Agri-Food Canada.

Agriculture and Agri-Food Canada. (2024). ISO 19131 – Annual Crop Inventory: Data Product Specifications (Revision A). Agroclimate, Geomatics and Earth Observation Division, Science and Technology Branch. https://open.canada.ca/data/en/dataset/ba2645d5-4458-414d-b196-6303ac06c1c9