

1

2

3

4

Abstract goes here

5

*Corresponding Author

Email addresses: samuel.robinson@ucalgary.ca (Samuel V. J. Robinson), hoanglan.nguyen@ucalgary.ca (Lan H. Nguyen), paul.galpern@ucalgary.ca (Paul Galpern)

1. Introduction

- Preserving SNL around fields is important for agriculture and conservation
 - Ecosystem services
 - Act as reservoirs for beneficial insects
 - Microclimate zones
 - Poorly studied in North America
 - Studies tend to be limited in scope (e.g. only certain organisms/crops) and applicability
- Ecosystem services can influence both the mean and variability of yield in agroecosystems
 - Typically only averages (means) are considered, but higher stability (lower variance) in yield is also valuable
- Field size and field boundaries are directly related to the conservation of SNL in agroecosystems
 - Size examples: wheat study from the UK
 - Boundary examples: flower strip studies, hedgerows
 - In North America, large financial incentives to make fields large and homogeneous, especially with large harvesting and planting equipment
 - Crop edge effects cause low yields at the margins of fields because of late emergence, poor microclimate, and competition with weeds
 - Ecosystem services should decay with distance from edge, so crops at the centre of a large field will not benefit from ecosystem services
 - Therefore, there should be a “goldilocks” field size, where negative edge effects are canceled out by ecosystem services
- Precision yield data holds enormous promise for agronomy
 - Limited because of:
 - Lack of standardized formats
 - Yearly calibration data
 - Clear statistical protocols

2. Methods

2.1. Data collection

- Precision yield data were collected directly from farmers across Alberta
 - Farmers were solicited for yield data through local agronomists, and we received data from a total of X growers across a total of X years
 - Most fields represented only a single year of data, but we did sometimes receive multiple years of data from the same field
 - File formats vary depending on the brand of combine, so we converted data to a standard format (csv) using Ag Leader SMS
 - In total, we analyzed yield data from X field-years of data, containing a total of X million data points
- Yield data was collected in discrete rectangles of the same length as the data interval (distance = speed \times interval, typically 1 second) and the same width as the combine header (5-7 m)
 - We extracted the size of each polygon (m^2), dry yield (tonnes), and the spatial location, and the sequence of collection (1 - end of harvest)
 - Because of the large number of yield rectangles per field (30-800 thousand), we used the centroid of each polygon as its location
 - Seeding and application rates were constant with each field, so we did not consider inputs in our analysis
- Field boundaries were automatically digitized, then manually checked using satellite imagery and classified land cover data
 - Crop boundaries are flexible, and often change from year to year depending on planting and emergence conditions (e.g. flooding during some years)
 - Additionally, seminatural features often change from year to year
 - * Ephemeral wetlands are flooded during some years, but consist mainly of grasses during dry years
 - * Grass boundaries can change if fields are used for as haying or pasture during crop rotation
 - This makes accurate and consistent classification of field boundaries very difficult

– Because of this, we used the following general categories for field boundaries:

1. Standard: thin (< 10 m wide) grassy field boundary, often grassy road right-of-ways
2. Wetland: permanent wetland, whose borders are largely unchanged from year-to-year
3. Forest/shrub: permanent windbreaks or remnant forests
4. Grass: larger grassy area (pasture or permanent seminatural grassland)
5. Other crop: different crop with little or no visible boundary between planted areas
6. Bare: unplanted (fallow), flooded area, or unpaved roads

2.2. Analysis

- At each field, we fit an additive model of the effect of boundary distance on crop yield while accounting for within-field spatial variation and temporal variation in the combine yield monitor
 - Crop yield varies within a field due to soil conditions, moisture, seeding rates, herbicide application, and previous agricultural practices such as strip farming
 - While sensor calibration can reduce combine-level bias (such as a combine recording consistently higher/lower yields), this does not address sensor drift over time that occurs within fields
 - The yield monitors may record lower yields as sensors accumulate debris during harvest (pers. comm. Trent Clark), leading to changes in accuracy and bias over time
 - Additionally, ground speed is known to be extremely important to yield monitor accuracy (Arslan & Colvin 2002)
 - To address this, we fit the following model:

$$\ln(yield) \tag{1}$$

3. Results

Results here

4. Discussion

Discussion here

85 **5. Authors' contributions**

86 Author's contribution

87 **6. Acknowledgements**

88 Funding for this research was provided by Ducks Unlimited Canada's Institute for Wetland and
89 Waterfowl Research, the Alberta Canola Producers Commission, Manitoba Canola Growers Association,
90 SaskCanola, the Alberta Biodiversity Monitoring Institute, and the Alberta Conservation Association.

91 **References**

92 Arslan, S. & Colvin, T.S. (2002). An evaluation of the response of yield monitors and combines to
93 varying yields. *Precision Agriculture*, 3, 107–122.

⁹⁴ **Appendix A: Supplementary Material**

⁹⁵ Supplemental materials here