United States
Department of
Agriculture

National
Agricultural
Statistics
Service

Research and
Development Division
Washington DC 20250

RDD Research Report
Number RDD-17-02

July 27, 2017

Comparing Geographic Information System (GIS) Calculated Acreage to Farmer Reported Acreage Utilizing a Mobile Mapping Instrument

Claire G. Boryan
Linda A. Lawson
Michael Hyman
Denise A. Abreu
Michael W. Gerling
Rick Hardin

EXECUTIVE SUMMARY

The United States Department of Agriculture's National Agricultural Statistics Service (NASS) conducts the annual June Area Survey (JAS) to obtain information on U.S. crops; livestock; grain storage capacity; and number, type and size of farms. During the JAS, field enumerators conduct personal interviews of farmers to obtain information on all agricultural activity occurring within a one-square-mile (approximately) sample unit or segment. Field enumerators are provided paper questionnaires and aerial photography that display the JAS segment boundaries. Field enumerators draw field boundaries on the aerial photo and use the paper questionnaire to record a farmer's agricultural activity within the segment and additional information about their entire operation.

June Area Survey segment boundaries are based on physical features (roads, railroads, rivers) and field enumerators use these physical features to identify segment locations. As a cost saving initiative, NASS is evaluating the use of 1) a permanent grid area sampling frame, with units of roughly equal size and area (one-square-mile) and 2) a prototype mobile mapping instrument for JAS data collection. This permanent grid sampling frame could be developed based on the Public Land Survey System's (PLSS's) one-square-mile sections in the 30 states in which the PLSS is the primary surveying method. The sample units of the permanent grid sampling frame are referred to as grid cells in this paper. In the 30 PLSS states, roads are often closely aligned with PLSS section lines. There are, however, exceptions to this rule and gaps in PLSS coverage exist. In states where land surveying is based on alternate systems (non-PLSS states), a grid sampling frame, with approximately one-square-mile grid cells, which are not based on physical features, could be generated using ESRI's ArcGIS software.

Due to the characteristics described above and the fact that both PLSS section lines and grid cells commonly cut across fields, a prototype mobile mapping instrument was developed to use Geographic Information Systems (GIS) technology for JAS data collection. In 2012, a team of researchers from NASS and Iowa State University's Center for Survey Statistics and Methodology developed a prototype mobile mapping instrument called Geographic Information Running Area Frame Forms Electronically (GIRAFFE). The instrument was designed to operate on an iPad and could be utilized to collect data for either grid cells or JAS segments.

The primary objective of this research was to determine whether the GIS calculation of acreage, based on fields digitized in a mobile mapping instrument, would result in field acreages that were comparable to acreages reported by JAS farmers. A permanent frame, based on grid cells for the JAS, would require the use of a mobile mapping instrument for data collection. If no significant difference was identified between the GIS calculation of acreage and JAS farmer reported acreages, research into the use of a permanent frame based on grid cells could proceed. Additionally, research into the use of a mobile mapping instrument for JAS data collection could advance with increased reliance on the GIS calculation of acreage

To compare the GIS calculation of acreage with JAS farmer reported acreage, a sample of 90 segments was selected in Indiana, Pennsylvania and Washington (30 segments per state). Indiana

and Washington are PLSS states and Pennsylvania is a non-PLSS state. Field enumerators were provided the completed aerial photos from the 2013 JAS and tasked with replicating field boundaries within the prototype mobile mapping instrument. The calculated GIS acreage was then compared with the acres reported by JAS farmers. Acreage comparisons were based on segment totals and at the field level for all crop fields, interior crop fields (< 25% of a field's border was also the segment border), and exterior or boundary crop fields (> 25% of a field's border was also the segment border).

To determine whether there was bias in the acreage of fields calculated using the GIS software, the median difference in acreage was examined and found to be zero. Wilcoxon signed-rank tests were performed for segment totals and at the crop field level for the three crop field categories. Although the results indicate that differences in the GIS calculation of acreage and JAS farmer reported acreage exist, *p*-values from the Wilcoxon signed-rank tests showed no evidence that median acreage differences were not equal to zero at the 95% confidence level, with the exception of exterior fields in Indiana. Because the distribution of acreage differences are centered at zero, show little bias, and have acceptable standard deviations, the authors believe that these differences between the GIS calculation of acreage and JAS farmer reported acreage will have a negligible impact on the estimates for the JAS.

RECOMMENDATIONS

1. Continue development of a prototype mobile mapping instrument, in an effort to modernize JAS data collection and reduce costs, with increased reliance on the GIS calculation of acreage

2. Study the feasibility of grid cell data collection using the prototype mobile mapping instrument, by conducting interviews with farmers in 2014

ACKNOWLEDGEMENTS

The authors would like to thank:

National Agricultural Statistics Service

Research and Development

Wendy Barboza

Jeff Bailey

Matt Deaton

Irwin Anolik

Field Operations

Debbie Dunham

Greg Matli

Kevin Pautler

Eric Stebbins

Jillayne Weaber

Indiana, Pennsylvania and Washington Field enumerators

Iowa State University

Center for Survey Statistics and Methodology

Alan Dotts

Eric Henely

Andrew Vardeman

TABLE OF CONTENTS

Comparing Geographic Information System (GIS) Calculated Acreage to Farmer Reported Acreage Utilizing a Mobile Mapping Instrument

Claire G. Boryan, Linda A. Lawson, Michael Hyman, Denise A. Abreu, Michael Gerling, and Rick Hardin

Abstract

The National Agricultural Statistics Service (NASS) conducts the annual June Area Survey (JAS), which is based on an area sampling frame. Segments of land comprise the JAS sampling units. Field enumerators use hard copy aerial photos that contain outlines of the sample unit or segment boundaries to locate and to interview all farmers within each sampled unit. Paper questionnaires are used to record the farmer's agricultural activity occurring within each segment and additional information about their entire operation.

June Area Survey sampled segment boundaries follow physical features (roads, railroads, rivers) on the ground. As a cost saving initiative, NASS is evaluating the use of a permanent grid area frame with sampling units of roughly equal size and area with data collection conducted using a mobile mapping instrument. The permanent grid frame would be based on the Public Land Survey System's (PLSS's) one-square-mile sections in the 30 states in which the PLSS is the primary surveying method. The sample units of the permanent grid frame are referred to as grid cells in this paper. In these 30 states, roads are often aligned with the PLSS section lines. However, exceptions to this rule and gaps in PLSS coverage exist. In states where land surveying is based on alternate systems (non-PLSS states), a grid frame, with one-square-mile sections would be generated using ESRI's ArcGIS software. These grid cells would not be aligned with physical features on the ground. Due to these characteristics and the fact that grids cells commonly cut across agricultural fields, a mobile mapping instrument was developed to use Geographic Information Systems (GIS) technology to calculate the acreage of the fields located in the grid cells.

The primary objective of this study is to determine whether a GIS calculation of acreage is comparable to the acreage reported by JAS farmers. If so, research into the use of a permanent grid sampling frame can move forward. Additionally, research into the use of a mobile mapping instrument for JAS data collection, based on grid cells or JAS segments, can advance with increased reliance on the GIS calculation of acreage. For this study, field enumerators used the aerial photos from previously collected 2013 JAS data to delineate field boundaries in the prototype mobile mapping instrument. The calculated GIS acreage, at the segment and crop field level, was then compared to the acreage reported by JAS 2013 farmers. Ninety segments in Indiana, Pennsylvania and Washington were included to assess segment and crop field acreage differences. Summary statistics and *p*-values from Wilcoxon signed-rank tests for total segment and crop field acreage differences between GIS calculated values and JAS farmer reported values are reported.

To determine whether there was any bias in segment-level and crop-field-level acreage, estimated using the GIS software, the median difference in acreage was examined and was found to be zero. Wilcoxon signed-rank tests were performed for segment totals and at field level for three crop field categories. Results indicated that, for segment totals and for all crop field types in all states, with

the exception of Indiana's boundary fields (p-value = 0.04), the GIS calculation of acreage and the JAS farmer reported acreages were not statistically different.

Key Words: Mobile Mapping, GIS Calculated Acreage, Data Collection, Area Frame Survey

1.0 INTRODUCTION

The National Agricultural Statistics Service (NASS) conducts the June Area Survey (JAS), which obtains information on U.S. crops; livestock; grain storage capacity; and the number, type and size of farms. During the annual JAS, field enumerators conduct personal interviews with farmers to obtain information on all agricultural activity occurring within a sample unit or segment that is approximately one-square-mile. Field enumerators are provided a hard copy aerial photo that displays the sampled segment boundaries, and a paper questionnaire. The field enumerators draw the field boundaries on the aerial photo and use the paper questionnaire to record information on the farmer's agricultural activity within the segment and additional information about the farmer's entire operation (Cotter *et al*., 2010; Ford, *et al*., 1986; Vogel, F., 1995; Nusser and House, 2009; Arroway *et al*., 2010).

June Area Survey segment boundaries are based on physical features (roads, railroads, rivers) and the field enumerators use these physical features to identify segment locations. As a cost saving initiative, NASS is evaluating the use of a permanent grid area frame, with sampling units of roughly equal size and area and a mobile mapping instrument for JAS data collection. The permanent frame would be based on the Public Land Survey System's (PLSS's) one-square-mile sections in the 30 states in which the PLSS is the primary surveying method (USGS, 2016). The sample units of the permanent grid frame are referred to as grid cells (Figure 4). In these 30 states, roads are often aligned with the PLSS section lines. In states, where land surveying is based on alternate systems (non-PLSS states), a grid sampling frame, with one square mile sections would be generated using ESRI's ArcGIS software. Figure 1 illustrates the grid sampling frame concept (outlined in yellow) compared with a JAS segment based on physical features (outlined in red).

Figure 1: A grid sampling frame with grid cells (outlined in yellow) compared with a JAS segment based on physical features (outlined in red)

Due to the general characteristics of the PLSS and the fact that PLSS section lines and grid cells commonly cut across fields, a mobile mapping instrument was developed to use Geographic Information Systems (GIS) technology for JAS data collection and specifically to calculate the acreage of the fields surveyed. In 2012, a team of researchers from NASS and Iowa State University's Center for Survey Statistics and Methodology developed a prototype mobile mapping instrument called Geographic Information Running Area Frame Forms Electronically (GIRAFFE) (Gerling *et al*., 2015). The instrument was designed to operate on an iPad and could be utilized to collect data for either grid cells or JAS segments.

An important question to answer before NASS could adopt the use of a mobile mapping instrument, based on GIS technology, was to determine if there was a statistically significant difference in the GIS calculation of acreage and JAS farmer reported acreage. Although GIS technology has advanced significantly in recent years, the JAS farmer reported acreages remain the basis for the NASS JAS crop estimates. Consequently, understanding if there is a statistically significant difference in the GIS calculation of acreage and JAS farmer reported acreage is foundational to the modernization of both the NASS area sampling frame and JAS data collection procedures.

GIS technology is currently used to calculate acreage within the USDA NASS and the Farm Service Agency (FSA). For example, NASS currently utilizes the GIS calculation of acreage within JAS segment boundaries as the basis for a 10% tolerance rule in which the JAS farmer reported acreages must sum to within 10% of the acreage calculated using the GIS software (USDA NASS, 2016). The USDA FSA relies upon the GIS calculation of acreage for many aspects of their acreage and compliance procedures (USDA FSA, 2016-a). GIS calculations of acreage are the foundation of the FSA Common Land Unit (CLU) program used to obtain crop acreage for FSA fields (USDA FSA, 2016-b). Historically, FSA calculated acreage on hard copy aerial photos using planimeters. However, FSA has transitioned to a comprehensive geospatial basis for the FSA farm records, which utilizes GIS software and aerial photography. Acreage for all CLU polygons is determined by digitizing around managed areas and field boundaries that are displayed over National Agricultural Imagery Program (NAIP) aerial photography. The FSA relies on farmer reports of crop acreage for CLU subfields but the CLU acreage calculation is made using GIS software. In letters sent to farmers notifying them of acreage changes and included as a reference in the FSA Handbook on Acreage and Compliance Procedures the " FSA reports that their tests have shown that the acreage calculation using computer software are equal to or slightly more accurate than the methodology used in the past" (USDA FSA, 2016-a). Other GIS technology, including Geographic Positioning System (GPS), is used to support staking and referencing procedures. GPS points are loaded into ArcGIS software and acreage calculations computed. The FSA recommends that calculated or measured acreage can be entered or maintained to "tenths or hundredths for higher precision and mitigation of rounding problems that may occur between acreage on FSA-578 and acreage in Farm Records". This high level of precision in the acreage calculation infers complete reliance upon the GIS calculation of acreage by the FSA (USDA FSA, 2016-a).

Based on the common usage of the GIS calculation of acreage both within NASS, and the FSA for acreage and compliance procedures, a statistical comparison of the GIS calculation of acreage and the JAS farmer reported acreage is warranted. The primary objective of this study is to determine whether a GIS calculation of acreage, using the mobile mapping instrument, is comparable to the acreage reported by JAS farmers. For this study, a sample of 90 segments was selected from the 2013 JAS conducted in Indiana, Pennsylvania and Washington (30 segments per state).

This paper is organized into nine sections. Section 2.0 includes background on the NASS Area Sampling Frame and the JAS. Section 3.0 describes the proposed permanent grid area sampling frame research. Section 4.0 describes the study areas. Section 5.0 describes the method used to conduct the assessments including 1) replication of JAS segments and field boundaries in the mobile mapping instrument; 2) the GIS calculation of acreage; and 3) the statistical tests. Section 6.0 includes the research results for segment totals and at the crop field level for each state individually and for all three states combined. Crop field level results are reported for 1) all crop fields, 2) interior fields, and 3) exterior or boundary fields. Section 7.0 includes a discussion of the research results and section 8.0 includes the conclusion. Section 9.0 describes recommendations for future research.

2.0 THE NASS AREA SAMPLING FRAME AND JUNE AREA SURVEY (JAS)

The NASS JAS is conducted annually utilizing an area sampling frame, which ensures complete coverage of all land in the U.S. Land within the area frame is divided into homogeneous strata based on percent cultivated land and further into substrata based on similarity of agricultural content. Within each stratum, the land is divided into primary sampling units (PSUs). A sample of PSUs is selected within substrata and smaller, similar-sized segments of land (approximately one- square-mile) are delineated within the selected PSUs. One segment is randomly selected from each selected PSU to be fully enumerated during the JAS. Selected segments usually have boundaries that follow physical features (roads, railroads and rivers) (Cotter, *et al.*, 2010).

June Area Survey segments (outlined in red in Figure 2) are pre-screened during the month of May prior to the June data collection period. During pre-screening, field enumerators divide each segment into separate tracts of land that represent each unique farm operating arrangement. Each tract is assigned a letter and drawn in blue on the aerial photo (Figure 2). Tracts are then screened for agricultural activity of which about 42,000 of them are classified as agricultural tracts. June Area Survey data collection is conducted during the first two weeks of June. At this time, field enumerators return to only the agricultural tracts and conduct interviews using the JAS questionnaire, which collects detailed agricultural information about the farmer’s land, both inside and outside the segment. Field enumerators complete a separate paper questionnaire for each agricultural tract operation within the segment. Farmers identify all field boundaries (outlined in red in Figure 3) on the aerial photo and report acreage and the crop planted or other land use of each individual field (pasture, woods, wasteland, etc.) (Cotter *et al*., 2010).

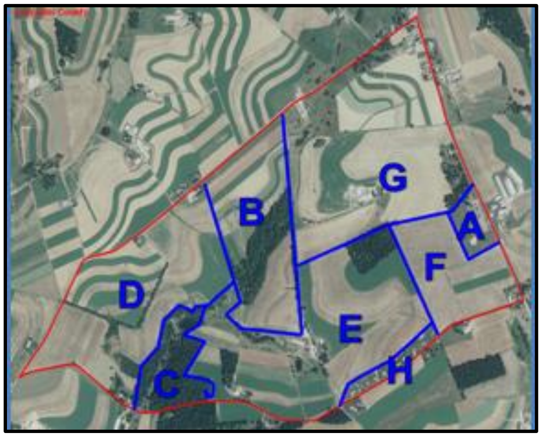

Figure 2: The segment is outlined in red. Tracts are outlined in blue and labeled with letters

Figure 3: Tracts are outlined in blue and labeled. Individual fields are outlined in red within the tracts and labeled with numbers

3.0 PERMANENT GRID SAMPING FRAME

NASS is evaluating the potential use of a permanent grid area sampling frame to replace the area sampling frame and segments based on physical features, as a cost saving initiative. The permanent grid frame sample units are roughly equal size and shape (approximately one-square-mile) and are referred to as grid cells in this paper. The permanent grid sampling frame could be based on the PLSS as described in Section 1.0 Introduction (USGS, 2016). In Figure 4, the 30 states included in the PLSS are identified in blue. Figure 5 illustrates PLSS section lines over a NAIP aerial photo (USDA FSA, 2016c). In Figure 5, the PLSS section lines are closely aligned with physical features on the ground. However, even in PLSS states, section lines do not perfectly coincide with roads, railroads or rivers, causing fields to be split. Further, in non-PLSS states, the grid cell boundaries are not related to physical features on the ground.

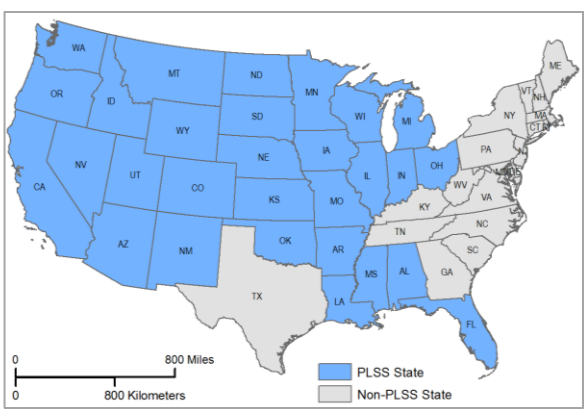

Figure 4: The thirty states included in the PLSS are highlighted in blue

Figure 5: PLSS section lines (red) displayed over NAIP aerial photography

Because grid cells commonly do not coincide with physical features and cut across crop fields, a mobile mapping instrument was developed to calculate field acreages using GIS technology. In 2012, NASS and Iowa State University's Center for Survey Statistics and Methodology researchers developed GIRAFFE, a prototype mobile mapping instrument. Gerling *et al*. provide detailed information about the instrument's functionality. The prototype mobile mapping instrument was designed to operate on an iPad and can be utilized to collect data for either grid cells or JAS segments (Gerling *et al.,* 2015). Figure 6 illustrates a grid cell digitized in the prototype mobile mapping instrument.

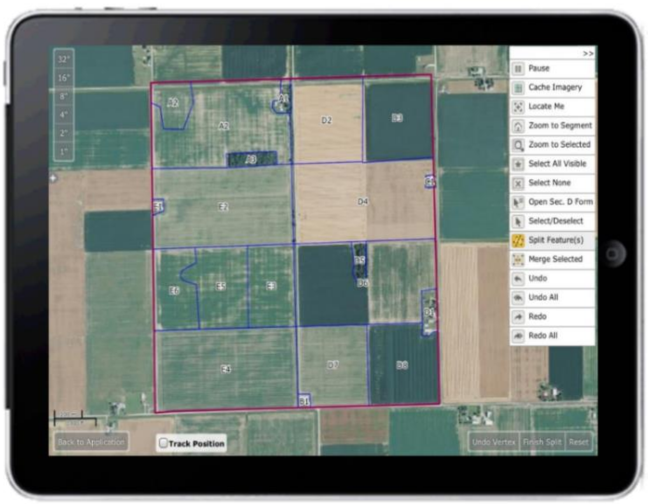

Figure 6: Prototype Mobile Mapping Instrument with a grid cell (red) displayed over NAIP aerial photography. The GIS tools are displayed are on the right side of the instrument.

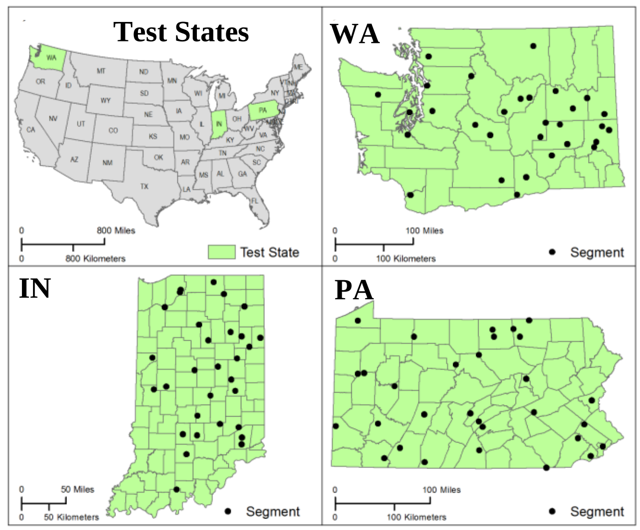

Figure 7: Study areas include: Indiana (IN), Pennsylvania (PA), and Washington (WA). Segment locations are identified by the black dots.

4.0 STUDY AREAS

Prior to determining whether the permanent grid sampling frame based on the PLSS and/or a mobile mapping instrument can be used within NASS, it was necessary to determine whether acreage information collected using a mobile mapping instrument and calculated using GIS software was comparable to acreage reported by farmers within JAS segments. Three states were selected as the study areas for this assessment. Indiana (IN) and Washington (WA) were selected as states with land surveyed using the PLSS system (Figure 4) and Pennsylvania (PA) was selected as an example of a state that was not surveyed using the PLSS system. The GIS calculation of acreage and the JAS farmer reported acreages were compared for 90 segments in IN, PA and WA (30 segments per state). The 90 segments were randomly selected from the 2013 JAS. Figure 7 illustrates the locations of the states and segments included in this assessment.

5.0 METHOD

This section describes the method used to conduct the assessments including 1) replication of JAS segment and field boundaries in the mobile mapping instrument; 2) the GIS calculation of acreage; and 3) the statistical analysis in which acreage comparisons were conducted using Wilcoxon signed-rank tests.

5.1 Replication of JAS tract and field boundaries in the mobile mapping instrument

The 90 segments selected for this research study were identified prior to the field enumeration stage of the 2013 JAS. Field enumerators were told that these 90 segments would be used for both the 2013 JAS and for the research described in this paper. The field enumerators were provided the instructions outlined in Appendix A. After the 2013 JAS was completed, the field enumerators, who participated in this study, were instructed to digitize the tract and field boundaries into the mobile mapping instrument (Figure 9) exactly as they were drawn on the 2013 JAS paper aerial photo (Figure 8). If a field enumerator conducted the original 2013 JAS interview for a segment, they were not assigned that segment.

Figure 8: A paper aerial photo identifying segment tract and field information collected during the 2013 JAS

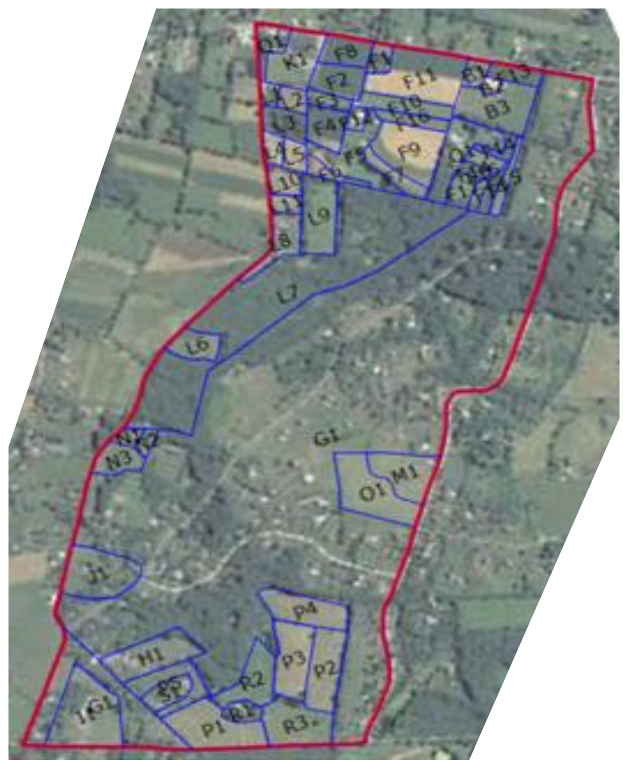

Figure 9: 2013 JAS segment in Figure 8 digitized in the mobile mapping instrument

5.2 GIS Calculation of Acreage

Once the field boundaries were digitized in the mobile mapping instrument, the corresponding field acreages were calculated directly from the polygon data (field boundaries) using ESRI's ArcGIS software. The polygons digitized in the mobile mapping instrument were imported into ArcGIS and the acreage calculated based on an appropriate projection and coordinate system. The GIS calculation of acreage was conducted using a customized Albers Equal Area Projection for the specific areas of interest. This is the same projection used in the construction of the NASS area frame and the selection of JAS segments.

A review of the GIS shapefiles, which included all field delineations, was conducted in order to identify potential problems or data anomalies. The only additional editing that occurred was the removal of four minor slivers of land. The slivers were not fields and were the result of an error that occurred when digitizing the field boundaries using the mobile mapping instrument. The slivers were merged into the appropriate adjacent fields. The sliver fields/polygons were so small (less than a tenth of an acre) that they had no effect on any of the acreage calculations (Figure 10).

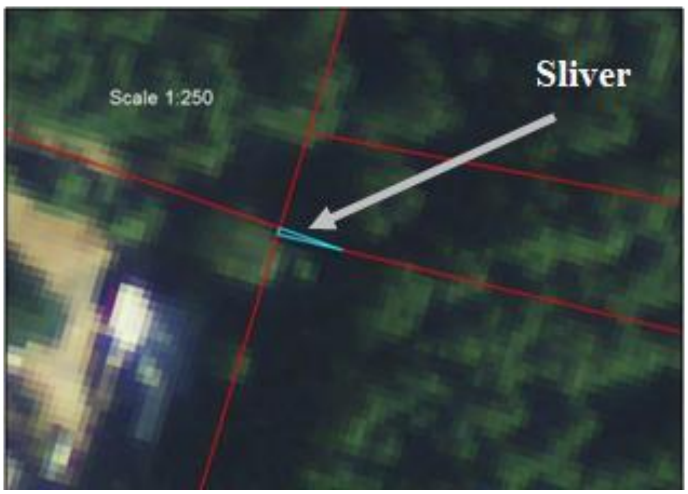

Figure 10: Sliver from a JAS segment identified in the mobile mapping instrument

5.3 Statistical Tests

An analysis was conducted to examine segment and crop field level differences between the GIS calculation of acreage and JAS farmer reported acreages. Acreage differences were determined by subtracting the JAS reported acreage for a segment or crop field from the acreage calculated using the GIS software. The acreage comparisons were conducted using Wilcoxon signed-rank tests to determine whether JAS farmer reported acreages for segments and crop fields were statistically significantly different from the GIS calculated acreages. The Wilcoxon signed-rank test is a nonparametric statistical test used to assess the null hypothesis that the distribution of a random variable is symmetric and centered at zero. The distributions that were tested included differences between GIS calculated acreage and JAS reported acreage for segments and three crop field categories. This statistical test was chosen due to non-normality of the acreage differences and small sample sizes for specific types of crop fields for which the sampling distribution was not assumed to be normal. Assumptions for the Wilcoxon signed-rank test are that observations are randomly selected and independent. In other words, the acreage differences between GIS and JAS acreages for a field or segment are independent from the acreage differences of other fields or segments. This assumption is valid for segment acreages because they do not share any boundaries.

The purpose of these comparisons was to assess whether the median acreage calculated using GIS software was statistically different from zero at the 95% confidence level when compared to the JAS farmer reported acreage values. Due to non-normality of the acreage differences, performing a statistical test to directly evaluate bias was difficult. The Wilcoxon signed-rank test evaluates whether the distribution of data is symmetrical and that the median of the distribution is equal to zero. The null hypothesis for these tests was that the median difference between GIS calculated acreage and JAS reported acreage for a segment is zero, versus the alternative hypothesis that the median acreage difference is not equal to zero. The tests were performed for segment acreage totals and for acreages at the crop field level, for each state individually and for the three states combined. Conclusions were drawn based on a 95% confidence or significance level of 0.05. The null hypothesis was rejected if the *p*-value was less than 0.05. Rejection of the null hypothesis was evidence that the distribution of acreage differences is not centered at zero. As sample size for the test increases, the mean of a symmetric distribution converges to the median. Thus, for large enough sample sizes, conclusions drawn from the Wilcoxon tests provide some indication of bias in field acreages calculated using a GIS. In addition to statistical tests, the mean, median, and standard deviation are reported as indications of whether acreages calculated using a GIS are biased and how much variation exists between the GIS calculated acreage and the JAS farmer reported acreage measurements.

The GIS calculation of acreage and the JAS farmer reported acreages were compared for agricultural tracts within 90 segments in IN, PA and WA (30 segments per state). Eight segments were removed from the analysis; three were considered outliers because they were composed of acreages much larger than other segments and were composed of only one, non-agricultural field. Five additional segments were removed due to errors in the JAS data collection (See Appendix B). Fields within each segment were digitized using the mobile mapping instrument and the field and segment acreages were compared to the acreages reported by farmers during the 2013 JAS. Tests regarding the distributions of differences in segment acreage totals consisted of the aggregated acreage differences over all fields within the segments. The final dataset consisted of 82 individual segments, which included 2,234 fields. These fields were further categorized based on the presence of cropland and the amount of border they shared with other fields and with the segment boundary.

6.0 RESULTS

This section includes the research results for segment totals and at the crop field level for each state individually and for all three states combined. Crop field level analysis was also conducted for 1) all crop fields, 2) interior fields and 3) exterior or boundary fields.

6.1 Segment Total Analysis

A total of 82 segments were included to assess segment total acreage differences including 27 segments in Indiana, 27 segments in Pennsylvania, and 28 segments in Washington (Table 1). To obtain segment total acreages for both GIS calculated and JAS farmer report values, field acreages were summed over all fields in a segment. Table 1 identifies the summary statistics and *p*-values from Wilcoxon signed-rank tests reporting total segment differences between the GIS calculation of acreage and JAS farmer reported acreage. These results are reported individually for each state and for all combined segments in the three states. A *p*-value of 0.05 or less indicates statistically significant differences at the 95% confidence level. Figure 11 displays histograms of the segment total acreage differences for each individual state and combined for all three states.

Table 1: Summary Statistics and *p*-Values for Wilcoxon Signed-Rank Tests for Segment Totals. All units of summary statistics are in acres.

State	Segments	Mean Difference	Standard Deviation	Median Difference	*p*-value
IN	27	-1.60	14.21	0.21	0.269
PA	27	4.64	26.02	1.30	0.835
WA	28	-0.89	35.14	-1.38	0.934
All Data	82	0.70	26.49	0.11	0.465

A *p*-value of 0.05 or less indicates significance at the 95% confidence level.

Mean differences between the GIS calculated segment acreage and JAS farmer reported segment acreage were less than five acres and the median acreage difference was less than one and a half acres for all three states. The mean and median acreage differences were 0.70 acres and 0.11 acres, respectively. Standard deviations of the segment level acreage differences ranged from 14.21 acres in Indiana to 35.14 acres in Washington. These results indicated that there were acreage discrepancies between the GIS calculation and JAS farmer reported acreage at the segment level. These individual segment level acreage discrepancies were largest in Washington. Individual acreage differences ranged from -74.2 to 107.1 acres for segments. The null hypothesis of the Wilcoxon signed-rank test was that the median difference between GIS calculated acreage and JAS reported acreage for a segment is zero, versus the alternative hypothesis that the median acreage difference was not equal to zero. The test results, with *p*-values of 0.269, 0.835 and 0.934 for segments in IN, PA and WA respectively, indicated that the median acreage difference (GIS minus JAS) was not statistically significantly different from zero for any of the three states.

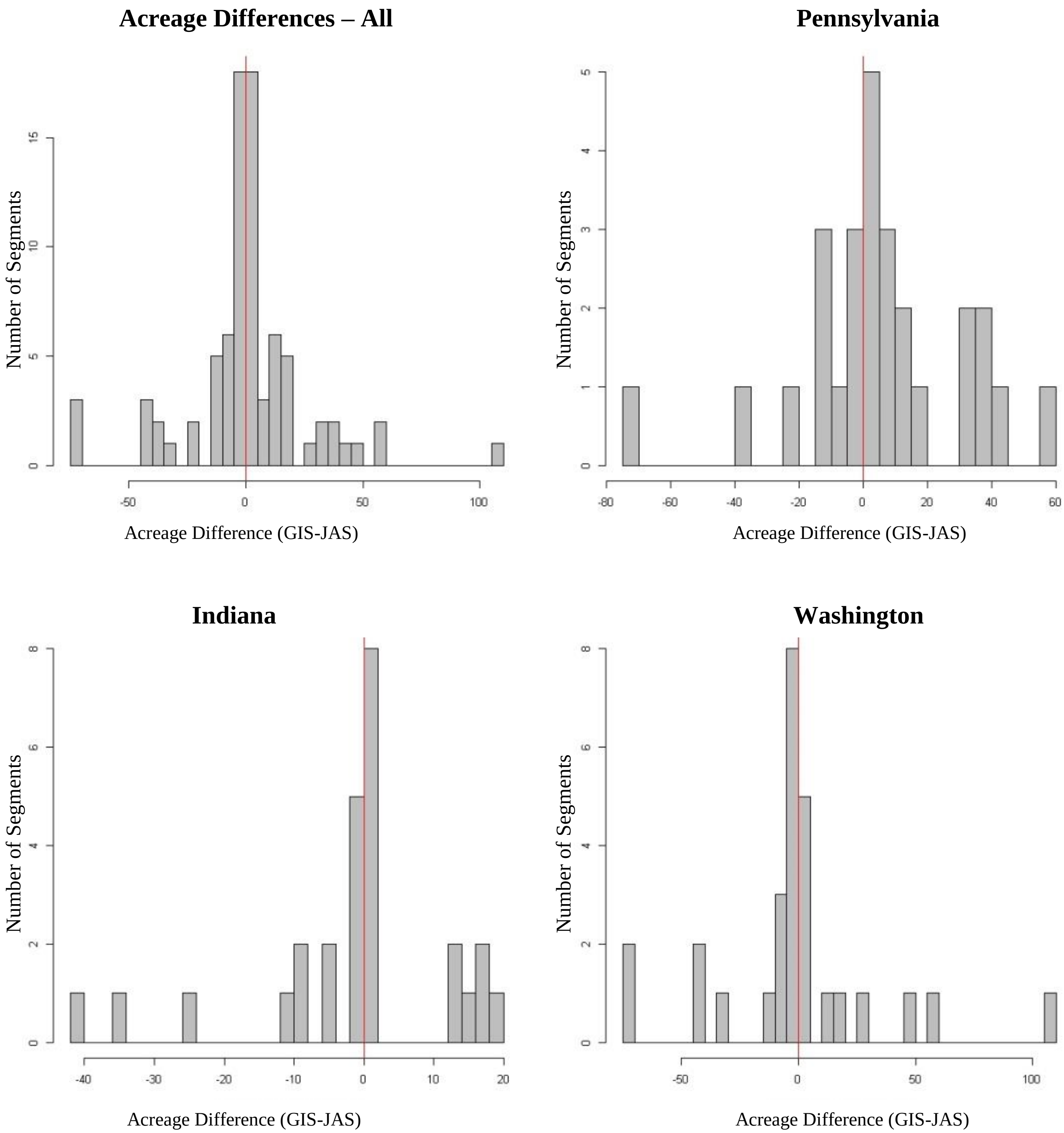

Figure 11: Acreage difference for all segments and at the segment level for each state individually. The red line indicates zero acreage difference.

6.2 Crop Field Analysis

A total of 2,234 fields composed the 82 segments in this analysis and 750 of these fields reported the presence of cropland on the 2013 JAS questionnaire: 351 in Indiana, 318 in Pennsylvania and 81 in Washington (Table 2). Approximately 90% of all crop fields were 100% cultivated, and no crop field in this analysis reported less than 50% cultivation during the 2013 JAS. Table 2 illustrates the summary statistics and *p*-values from Wilcoxon signed-rank tests of differences between the GIS calculation of acreage and JAS farmer reported acreages for all crop fields. Figure 12 displays histograms of the acreage differences for all crop fields and for each individual state.

Summary statistics showed that the mean and median differences between the GIS calculation of crop field acreage and the corresponding JAS farmer reported acreage for the same crop fields was less than 0.4 acres for all three individual states. Standard deviations of crop field acreage differences ranged from 3.72 acres in Pennsylvania to 24.18 acres in Washington. This indicated discrepancies between the GIS calculation and JAS farmer reported field acreage for individual crop fields and that these differences were larger in Washington. Individual acreage differences ranged from -99.43 to 66.99 acres for crop fields. However, the mean acreage difference for all crop fields was only 0.08 acres. The null hypothesis tested using a Wilcoxon signed-rank test was that the median difference between the GIS calculation and JAS farmer reported acreage for a crop field was equal to zero, versus the alternative hypothesis that the median acreage difference was not equal zero. Test results indicated that the median acreage difference (GIS minus JAS) was not statistically different from zero for any of the three states at the 5% significance level.

Table 2: Summary Statistics and *p*-Values for Wilcoxon Signed-Rank Test for All Crop Fields. All units of summary statistics are in acres.

State	Crop Fields	Mean Difference	Standard Deviation	Median Difference	*p*-value
IN	351	0.05	4.37	0.15	0.223
PA	318	0.22	3.72	0.07	0.668
WA	81	-0.32	24.18	0.40	0.706
All Data	750	0.08	8.79	0.15	0.184

A *p*-value of 0.05 or less indicates significance at the 95% confidence level.

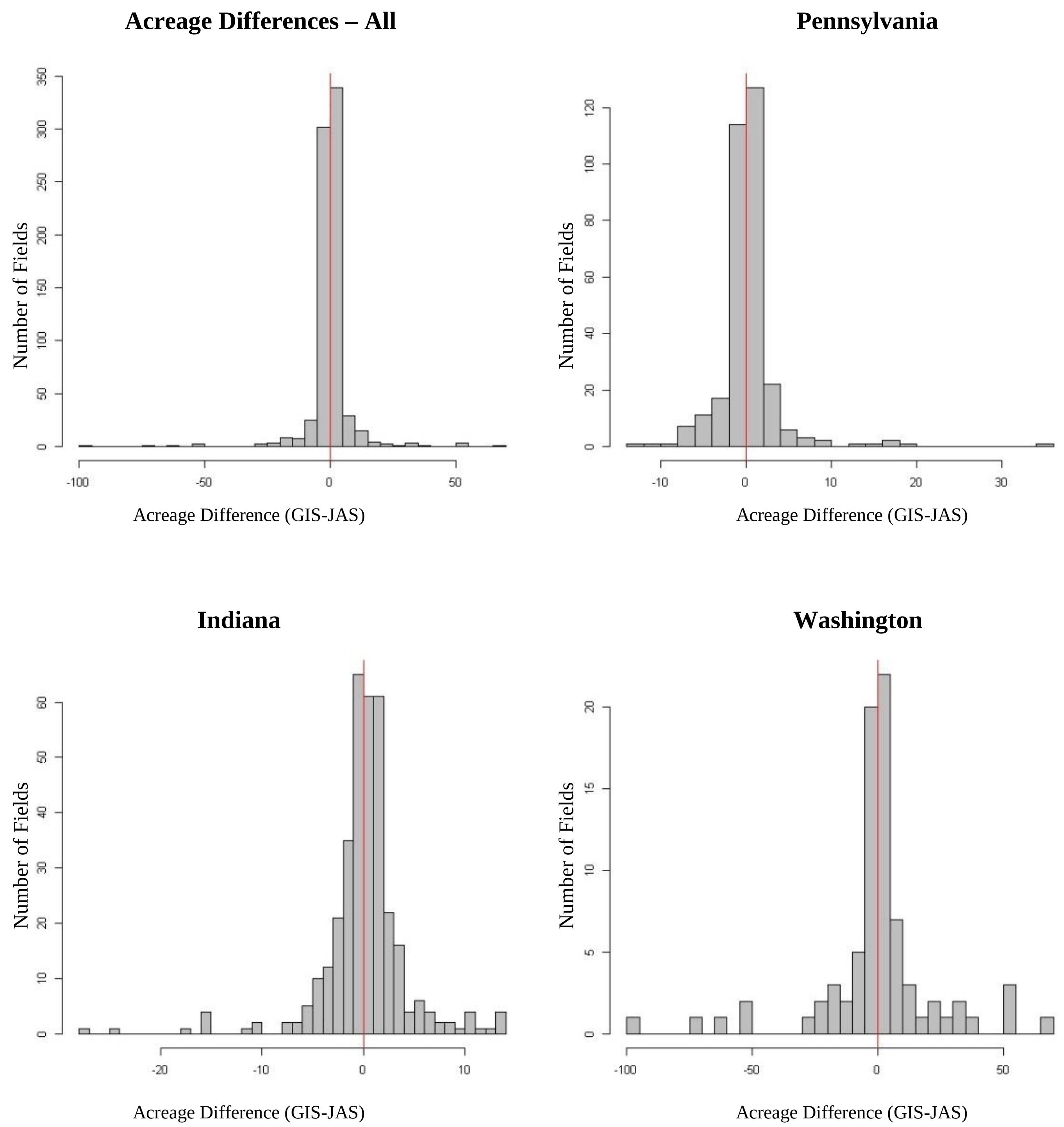

Figure 12: Acreage difference for all crop fields and crop fields for each state individually. The red line indicates zero acreage difference.

6.2.1 Interior Crop Fields

For this assessment, interior crop fields were defined as crop fields that had at least 75% of their perimeter bordering neighboring fields (i.e., less than 25% of the field's border was also the segment border). Thus, interior fields shared a border with at least one adjacent field. Particular interest lies in GIS calculations of acreage for interior fields because often, only a single border is digitized in the mobile mapping device separating the two adjacent fields. Both field's GIS acreages are determined by this border so there is a potential for correlation between neighboring fields' acreages. Interior fields share a larger percentage of their border with adjacent fields and have a higher potential for larger acreage differences. Thus, the distribution of the difference between the GIS calculation of acreage and JAS farmer reported acreage was examined. The segments used in this analysis contained 579 interior crop fields, including 259 in Indiana, 271 in Pennsylvania and 51 in Washington. Table 3 shows the summary statistics and *p*-values from Wilcoxon signed-rank tests for interior crop fields' differences between the GIS calculation and the JAS farmer reported acreages. These results are reported individually for each state and for all interior crop fields in the three state combined. Figure 13 displays histograms of the acreage differences for all interior crop fields for each individual state and combined for all three states.

For interior crop fields, mean differences between the GIS calculation and JAS farmer reported acreage ranged from -0.19 acres in Indiana to 0.7 acres in Washington. The overall mean acreage difference was 0.08 acres for all three states. The median acreage difference was closer to zero than the mean acreage difference for all three states, and was 0.06 acres for all interior crop fields. Indiana and Pennsylvania had little variation in field acreage differences, with standard deviations of 4.08 and 3.93 acres respectively. However, Washington had crop fields with large discrepancies between the GIS calculated acreage and JAS reported acreage, and had a standard deviation of 24.63 acres. The acreage difference of individual fields ranged from -99.43 acres to 66.99 acres in Washington. A Wilcoxon signed-rank test was used to test the null hypothesis that the median difference between the GIS calculation of acreage and JAS farmer reported acreage for interior crop fields was equal to zero, versus the alternative hypothesis that the median acreage difference was not equal to zero. The tests indicated that the median acreage difference (GIS minus JAS) was not statistically different from zero for any of the three states.

Table 3: Summary Statistics and *p*-Values for Wilcoxon Signed-Rank Test for Interior Agricultural Fields. All units of summary statistics are in acres.

State	Interior Crop Fields	Mean Difference	Standard Deviation	Median Difference	*p*-value
IN	259	-0.19	4.08	0.07	0.848
PA	271	0.22	3.93	0.01	0.971
WA	51	0.70	24.63	0.40	0.407
All Data	579	0.08	8.19	0.06	0.629

A *p*-value of 0.05 or less indicates significance at the 95% confidence level.

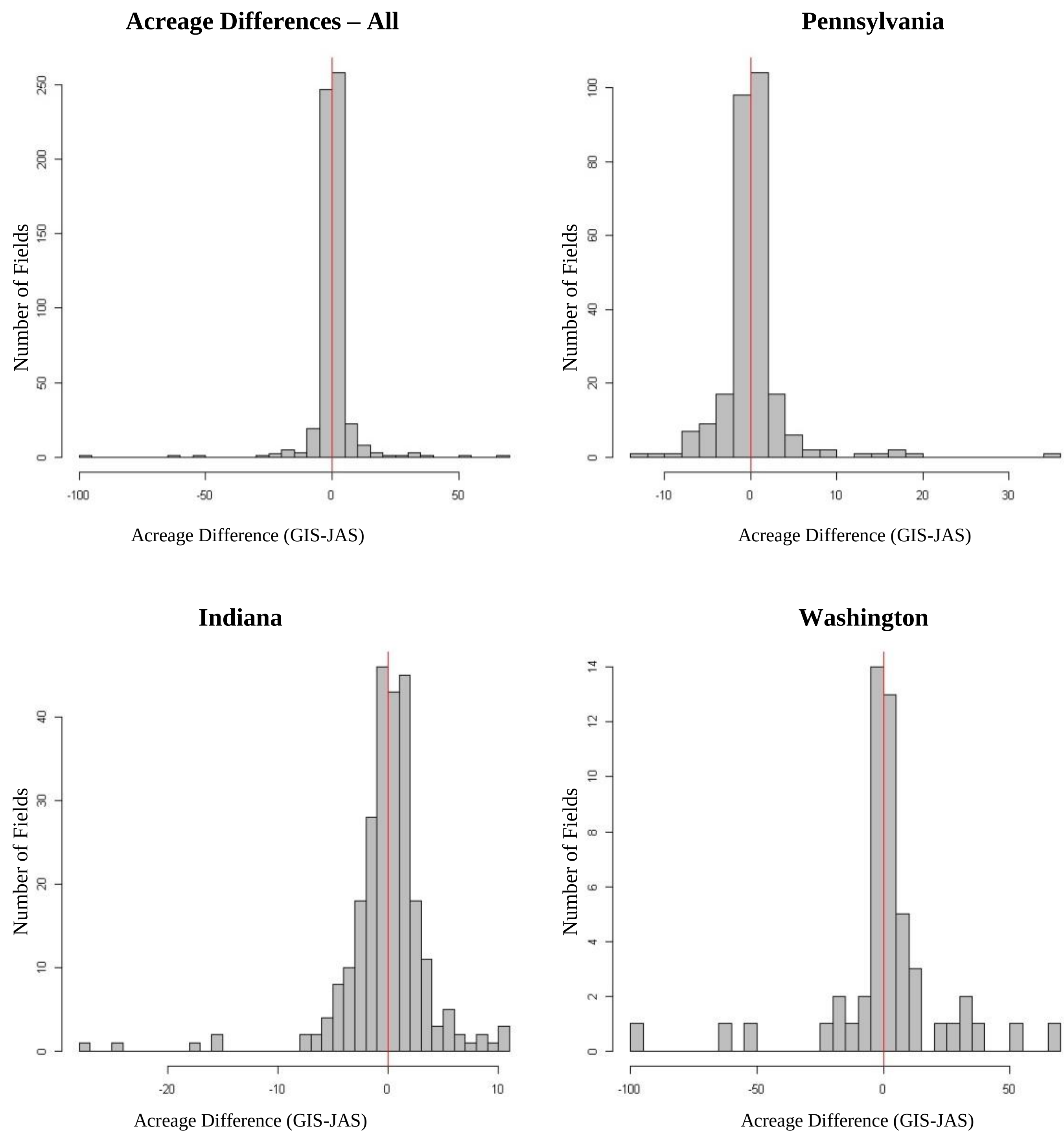

Figure 13: Acreage difference for all interior crop fields and for interior crop fields for each state individually. The red line indicates zero acreage difference.

6.2.2. Exterior or Boundary Crop Fields

The differences between the GIS calculation and JAS farmer reported acreage were also examined for exterior or boundary crop fields in the three test states. For this assessment, exterior or boundary crop fields were fields that had less than 75% of their perimeter bordering neighboring fields (i.e., at least 25% of the field's border was also the segment border) and also reported containing cropland during the 2013 JAS. Table 4 identifies the summary statistics and *p*-values from the Wilcoxon signed-rank tests comparing the GIS calculation and JAS farmer reported acreages for the exterior or boundary crop fields. The test results are reported individually for each state and for all exterior fields in the three states combined. Figure 14 displays histograms of the field acreage differences. The 82 segments used in this analysis contained a total of 71 exterior crop fields including 94 in Indiana, 47 in Pennsylvania and 30 in Washington (Table 4).

Table 4: Summary Statistics and *p*-Values for Wilcoxon Signed-Rank Test for Exterior Crop Fields.

State	Exterior Crop Fields	Mean Difference	Standard Deviation	Median Difference	*p*-value
IN	94	0.70	5.05	0.43	0.045*
PA	47	0.26	2.10	0.27	0.241
WA	30	-2.06	23.70	0.21	0.670
All Data	171	0.09	10.58	0.36	0.072

A *p*-value of 0.05 or less indicates significance at the 95% confidence level.
***Significant at the 0.05 level.**

Acreage differences in exterior crop fields were larger on average than for interior crop fields. The mean acreage differences for the three individual states ranged from -2.06 acres in WA to 0.26 acres in PA. However, the mean acreage difference for all exterior crop fields combined was only 0.09 acres. The median acreage differences were similar across individual states and was 0.36 acres for all combined exterior crop fields. Indiana and Pennsylvania had little variation in exterior crop field acreage differences, with standard deviations of 5.05 and 2.10 acres respectively. However, Washington had fields with larger discrepancies between the GIS calculated acreage and JAS reported acreage, and had a standard deviation of 23.70 acres. A Wilcoxon signed-rank test was used to test the null hypothesis that the median difference between the GIS calculation of acreage and JAS farmer reported acreage for exterior crop fields was equal to zero, versus the alternative hypothesis that the median acreage difference was not equal to zero. The tests indicated moderately strong evidence that the median acreage difference (GIS minus JAS) was statistically different from zero for Indiana and marginal evidence that the median was not equal to zero for all combined fields. The median acreage difference was not significantly different from zero for Pennsylvania and Washington.

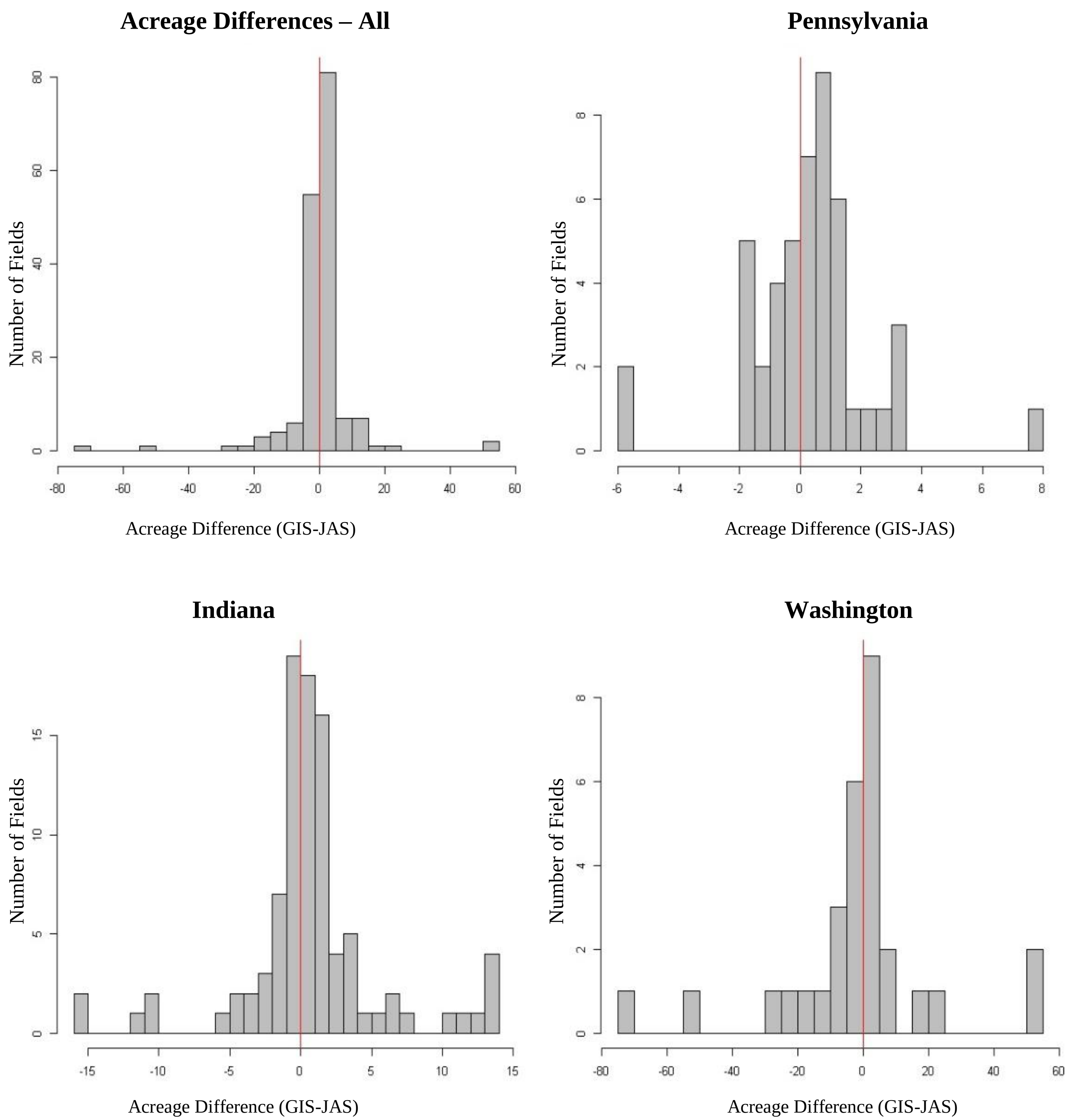

Figure 14: Acreage differences for all exterior or boundary fields and for exterior or boundary fields for each state individually. The red line indicates zero acreage difference.

7.0 DISCUSSION

Based on the analysis of eighty-two 2013 JAS segments in Indiana, Pennsylvania and Washington, segment level summary statistics showed that the mean acreage difference between the GIS calculated acreage and JAS farmer reported acreage at the segment level was less than five acres and the median acreage difference was less than one and a half acres for all three states examined. These were small differences considering that the average segment size was 640 acres. The mean and median acreage differences were 0.70 acres and 0.11 acres, respectively, for segment total acreages. Although discrepancies between the GIS calculation of acreage and the JAS farmer reported acreage existed at the segment level, the test results indicated that the median difference between GIS calculated acreage and the JAS farmer reported acreage was not significantly different from zero in any of the three states.

As was the case with the segment level acreage differences, the *p*-values from Wilcoxon signed-rank tests, which compared the GIS calculation of acreage and the JAS farmer reported acreage, indicated that the median difference between GIS calculated acreage and the JAS farmer reported acreage was not significantly different from zero for all crop fields. Interior crop fields and exterior or boundary crop field acreages were also compared. The interior crop field calculations of acreage based on the GIS and the JAS famer reported acreages had *p*-values of 0.848 in Indiana, 0.971 in Pennsylvania and 0.407 in Washington, indicating that the median difference in GIS calculations of acreage and JAS farmer reported acreages were not significantly different from zero.

The only crop field category with significant evidence that the median difference between the GIS calculated acreage and JAS farmer reported acreages was not equal to zero was for exterior or boundary fields in Indiana. The Wilcoxon signed-rank test, for exterior or boundary crop fields in Indiana had a *p*-value of 0.045. After reviewing the completed 2013 JAS aerial photos, the paper questionnaires, and the digitized field boundaries in the mobile mapping instrument; it was determined that the 2013 JAS field enumerators included trees, hedgerows, and drainage ditches along the edge of the segment boundary, and the JAS farmer operators reported only their tillable land. Consequently, when the field enumerators were instructed to digitize the field boundaries into the mobile mapping instrument exactly as they were drawn on the aerial photos, a larger total acreage (which included the trees, hedgerows and drainage ditches) was calculated using the GIS software. Figures 15 and 16 illustrate two fields in which the difference between the GIS calculation of acreage and JAS farmer reported acreage was greater than 4 acres. Figure 15 illustrates a crop field, digitized in the mobile mapping instrument, that includes a road that extends along the south and east side and a group of trees contained within the boundary on the west side. Figure 16 illustrates a crop field that includes a road extending along the south and east sides. Although the 2013 JAS field enumerators included these features within the boundaries of the crop fields, farmers were not likely to report this acreage during the JAS. Farmers generally report the acreage of tillable land or land that is useable for its purpose (cropland, pasture, etc.). Field enumerators used the JAS aerial photo field boundaries as a guideline for defining each field and thus followed them closely when digitizing fields in the mobile mapping instrument. This leads to greater acreage estimates when the GIS acreage is used as compared to the acreage reported by the farmers. This information is useful as field enumerators will need to be trained to identify these boundary features (roads, tree lines, hedgerows, etc.) and to not include them within the crop field boundary when using a mobile mapping instrument.

Figure 15: Exterior or boundary field in Indiana. This field includes a road that extends along the south and east side and a group of trees contained within the boundary on the west side.

Figure 16: Exterior or boundary field in Indiana. This field contains a road extending along the south and east sides.

8.0 CONCLUSION

The main objective of this research was to determine whether a GIS calculation of acreage was comparable to the acreage reported by JAS farmers. For this study, enumerators used the aerial photos from previously collected 2013 JAS data to digitize field boundaries in a prototype mobile mapping instrument. The calculated GIS acreage at the segment and crop field level was compared to the acres reported by JAS 2013 farmers. Eighty-two segments in Indiana, Pennsylvania and Washington were used to assess segment and crop field acreage differences. Summary statistics and *p*-values from Wilcoxon signed-rank tests for total segment and crop field acreage differences between GIS calculated values and JAS farmer reported values were reported. Test results indicated that all median acreages calculated at the segment and crop field level with the exception of exterior or boundary fields in Indiana had *p*-values less than 0.05 indicating that the median difference in acreage between the GIS calculation of acreage and JAS farmer reported acreages was not significantly different from zero.

Although differences in the GIS calculation of acreage and JAS farmer reported acreage exist, *p*-values from the Wilcoxon signed-rank tests showed no evidence that median acreage differences were not equal to zero at the 95% confidence level, with the exception of exterior fields in Indiana. Because the distribution of acreage differences are centered at zero, show little bias, and have acceptable standard deviations, the authors believe that these differences between the GIS calculation of acreage and JAS farmer reported acreage will have a negligible impact on the estimates for the JAS. Based on the results of this research, the authors recommend that NASS 1) continue development of a prototype mobile mapping instrument in an effort to modernize JAS data collection and reduce costs with increased reliance on a GIS calculation of acreage and 2) study the feasibility of grid cell data collection using the prototype mobile mapping instrument, by conducting interviews with farmers in 2014.

9.0 RECOMMENDATIONS

1. Continue development of a prototype mobile mapping instrument in an effort to modernize JAS data collection and reduce costs with increased reliance on a GIS calculation of acreage

2. Study the feasibility of grid cell data collection, using the prototype mobile mapping instrument, by conducting interviews with farmers in 2014

10.0 REFERENCES

Arroway, P., Abreu, D., Lamas, A., Lopiano, K., and Young, L. (2010). An alternate approach to assessing misclassification in JAS. In: Proceedings of the section on survey research methods JSM 2010, American Statistical Association, Alexandria VA

Cotter, J. Davies, C., Nealon, J., & Roberts R. (2010). *Area Frame Design for Agricultural Surveys in Agricultural Survey Methods* (eds R. Benedetti, M. Bee, G. Espa and F. Piersimoni), John Wiley & Sons, TD, Chichester, UK, doi:10.1002/9780470665480.ch11

FAO, "*Multiple Frame Agricultural Surveys. Volume 2* (1998): *Agricultural Survey Programmes based on Area Frame or Dual Frame (Area and List) Sample Designs*." FAO Statistical Development Series no. 10. Rome: FO. 274.

Ford, B. I., Nealon, J., Tortora, R. D (1986). Area frame estimators in agricultural surveys; sampling versus non sampling errors. Agricultural Economics research 38 (2), 1-10.

Gerling, M., Lawson, L., Weaber, J., Dotts, A., Vardeman, A., and Wilson E. (2015). NASS Report: Field Data Collection using Geographic Information Systems Technologies and iPads on the USDA's June Area Frame Survey.

Nusser S and House C. (2009) Sampling, data collection, and estimation in agricultural surveys. In: Pfeffermann D, Rao C (eds) Handbook of statistics 29A, sample surveys: design, methods and applications. Elsevier. The Netherlands, pp 471-486.

USDA FSA (2016-a). Farm Service Agency Handbook: Acreage and Compliance Determinations <https://www.fsa.usda.gov/Internet/FSA_File/2-cp_r15_a97.pdf last accessed 04/10/2016> (last accessed 04/12/17).

USDA FSA (2016-b) Farm Service Agency: Common Land Unit <https://www.fsa.usda.gov/programs-and-services/aerial-photography/imagery-products/common-land-unit-clu/index> (last accessed 04/12/17).

USDA FSA (2016c). National Agricultural Imagery Program (NAIP) Imagery, <http://www.fsa.usda.gov/programs-and-services/aerial-photography/imagery-programs/naip-imagery/> (last accessed 04/12/2017).

USDA NASS (2016). Survey Administration Manual (Internal Report) Section 6.2.3, Washington, D.C.

USGS (2016). The U.S. Map: The Public Land Survey System. <http://nationalmap.gov/small_scale/a_plss.html> (last accessed 04/12/17).

Vogel, F. A. (1995). The evolution and development of agricultural statistics at the United States Department of Agriculture. *Journal of Official Statistics*, 11: 161-180.

APPENDIX A

FIELD ENUMERATOR'S DATA COLLECTION FORM

State: ___________ **2013 Acreage Test Research Project** Segment: ___________

This segment has been randomly selected for a research project. This research will play an important role in reducing respondent burden in that it will help us determine if we need to continue to ask the acreage or if the new instrument's acreage calculation is close to the operator's response.

Part of the test will be to duplicate the field boundaries as drawn on the JAS aerial photos into the GIRAFFE instrument. Therefore, it is particularly important that grease pencil markings are legible, accurate, and precise in regard to field boundaries.

It would be extremely helpful to the research if you could record some additional information about how the acreage within section D was obtained during June. Please answer the following questions for each Ag-Tract:

1. Ag tract letter: ____
2. Who supplied the **field acres** recorded on line #1 in **Section D**? ☐ Operator/Manager ☐ Spouse
   ☐ Accountant/bookkeeper ☐ Partner ☐ Other ☐ Nobody, the enumerator observed & gridded it
3. How would you describe the respondent's ability to report the **field acreage**?
   ☐ Respondent easily reported the acres for **every field** without the enumerator's aid or using the acreage grid
   ☐ They were able to easily report the acres for just these field numbers: ______________________________
   ☐ Respondent **wasn't sure of the acreage** or refused to provide it for **any** of the fields.
4. Please record any notes that you think would be helpful when recreating the field boundaries in section D:

   ____

1. Ag tract letter: ____
2. Who supplied the **field acres** recorded on line #1 in **Section D**? ☐ Operator/Manager ☐ Spouse
   ☐ Accountant/bookkeeper ☐ Partner ☐ Other ☐ Nobody, the enumerator observed & gridded it
3. How would you describe the respondent's ability to report the **field acreage**?
   ☐ Respondent easily reported the acres for **every field** without the enumerator's aid or using the acreage grid
   ☐ They were able to easily report the acres for just these field numbers: ______________________________
   ☐ Respondent **wasn't sure of the acreage** or refused to provide it for **any** of the fields.
4. Please record any notes that you think would be helpful when recreating the field boundaries in section D:

   ____

APPENDIX B

EVALUATION OF ERRORS FOUND IN THE COMPARISON OF GIS VS. JAS ACREAGE

Table B1 displays the distribution of errors and acreage discrepancies by state.

Table B1: Errors found in comparison of JAS vs. GIS Acreage

Discrepancy	IN	PA	WA	Total
Non-matching fields error	10	20	102	132
Large acreage difference	44	67	56	167
Total	54	87	158	299

For this evaluation, scanned images of the 2013 JAS questionnaires and supplementary forms specific to the fields in question, were used to derive information on the respondent and how easily acreage information was provided. A visual comparison of segments in the mobile mapping instrument with the aerial photos determined whether boundaries were positioned correctly by the enumerator. For the aerial photo, an acreage grid was used to verify acreage provided by the JAS respondent (farmer). Finally, error categories were created and all identified fields were classified as one of the resulting categories.

Table B2 describes the non-matching fields in which acreage was not reported in the JAS or not estimated in the mobile mapping instrument. The vast majority of these types of errors occurred in WA and resulted from other types of errors associated with the JAS.

Table B2: Analysis of Reasons for Non-matches between JAS and GIS Acreage

Reason for Non-matching Field	IN	PA	WA	Total
JAS labeling procedures	0	0	41	41
JAS enumerator mislabeled field	4	2	8	14
Other errors associated with JAS	4	15	50	69
Mobile mapping instrument enumerator mislabeled field	2	2	0	4
Other errors associated with using the mobile mapping instrument	0	1	3	4
Total	10	20	102	132

Table B3 describes fields with large acreage differences (> 10 acres difference) in which errors were identified prior to the analysis. The majority of the large acreage differences were a result of enumerator gridding error. This finding was consistent across all states.

Table B3: Errors associated with JAS data collection in fields with large acreage (>10 acres) differences between JAS and GIS Acreage

Reason for Acreage Difference	IN	PA	WA	Total
Enumerator gridding error	19	40	31	90
Incorrect field boundary position	6	3	5	14
Incorrect reported acreage	6	12	0	18
Other errors associated with JAS	4	9	15	28
Total	35	64	51	150

Table B4 describes the errors associated with the fields with large acreage differences between JAS and mobile mapping instrument estimates that were not due to either enumerator or farmer error. These errors are associated with the mobile mapping instrument GIS acreage calculation process.

Table B4: Errors Associated with the Mobile Mapping Instrument Occurring in Fields with Large Acreage (>10 Acres) Differences between JAS vs GIS Acreage

Reason for Acreage Difference	IN	PA	WA	Total
Precise boundary constraint	6	0	4	10
Incorrect field boundary position	3	3	1	7
Total	9	3	5	17