Article

Estimates of Conservation Tillage Practices Using Landsat Archive

Peter C. Beeson [1,*]**, Craig S.T. Daughtry** [2] **and Steven A. Wallander** [3]

[1]   Global Conservation Institute, Santa Fe, NM 87506, USA
[2]   Agriculture Research Service, USDA, Beltsville, MD 20705, USA; craig.daughtry@usda.gov
[3]   Economic Research Service, USDA, Washington, DC 20250, USA; steve.wallander@usda.gov
[*]   Correspondence: pbeeson@global-conservation.org

check for updates

Abstract: The USDA Environmental Quality Incentives Program (EQIP) provides financial assistance to encourage producers to adopt conservation practices. Historically, one of the most common practices is conservation tillage, primarily the use of no-till planting. The objectives of this research were to determine crop residue using remote sensing, an indicator of tillage intensity, without using training data and examine its performance at the field level. The Landsat Thematic Mapper Series platforms can provide global temporal and spatial coverage beginning in the mid-1980s. In this study, we used the Normalized Difference Tillage Index (NDTI), which has proved to be robust and accurate in studies built upon training datasets. We completed 10 years of residue maps for the 150,000 km^2 study area in South Dakota, North Dakota, and Minnesota and validated the results against field-level survey data. The overall accuracy was between 64% and 78% with additional improvement when survey points with suspect geolocation and satellite tillage estimates with fewer than four dates of Landsat images were excluded. This study demonstrates that, with Landsat Archive available at no cost, researchers can implement retrospective, untrained estimates of conservation tillage with sufficient accuracy for some applications.

Keywords: crop residue cover; plant litter; non-photosynthetic vegetation; soil disturbance; normalized difference tillage index (NDTI); environmental quality incentives program (EQIP)

1. Introduction

The USDA Environmental Quality Incentives Program (EQIP) promotes the use of conservation practices by providing financial and technical assistance to producers. One important question for analysis of the program is what proportion of producers continue those practices without payment once the contracts end. The challenge in analyzing this and similar questions related to program evaluation is the lack of adequate long-term data on the adoption of these conservation practices.

Research to address these issues requires field-level data for farms that are enrolled in EQIP conservation contracts as well as field-level data for farms that are not enrolled. Data collection is difficult because farming operations—e.g., tilling, planting, harvesting—are done over a range of times. The USDA-National Agricultural Statistics Service Crop Progress and Condition Reports indicate planting progress and phenological development for each crop at the state scale, but miss the mosaic of planting dates for individual fields at the farm scale (https://quickstats.nass.usda.gov/). One practice of interest is the adoption of conservation tillage practices, which increase the proportion of the soil surface covered by crop residue (or litter). Crop residue cover is typically assessed shortly after spring planting operations have been completed, where <15% residue cover is classified as intensive or conventional tillage, 15% to 30% residue cover is classified as reduced tillage, and residue >30% is classified as conservation tillage [1]. Reducing tillage intensity enhances numerous ecosystem services including

increased soil organic matter, decreased soil erosion, improved soil and water quality, and reduced carbon dioxide emissions [2].

Minimum crop residue cover typically occurs shortly after planting, but planting dates within a state often vary by more than two months due to a wide range of reasons because different crops require different planting dates (e.g., corn precedes soybean by weeks), and days unsuitable for fieldwork (e.g., soils may be too cold or too wet) can delay planting. While satellite imagery can be useful in determining these on a field scale [3–7], it can be difficult for many reasons. For images in the visible and infrared range, weather can render it useless. If a single well-timed, cloud-free satellite image exists, it might capture the variation, but it is likely that some fields are either not yet planted or planted too long ago. Minimum crop residue cover can be missed if the field is not yet planted, and in fields whose crop has already started to grow, the appearance of the soil/residue may be obscured. To overcome this problem, one merely needs to have more images with the hope that some will be cloud-free; but at a 16-day overpass schedule, there might be as many as 32 days between images or worse if clouds persist on overpass days. This can be improved by using multiple satellites and utilizing the overlap area between paths to sometimes have as little as one day between images, but only for a small portion of the land. Using different satellite platforms can reduce the overpass time in half for every satellite added.

There are many index-derived methods for estimating crop residue cover using satellites. Although hyperspectral imagery can produce higher accuracy crop residue cover estimates [3,8–10], it does not have the temporal recurrence interval or spatial coverage to capture the field-level minimum residue cover seen at different times throughout the spring. New commercial satellites, such as WorldView-3, continue to improve the band and pixel size requirements for better field-scale tillage estimates, but they have limited spatial coverage [11]. To best capture the landscape mosaic seen in crop residue cover in the landscape, a compilation of images is required [12,13]. The Landsat Thematic Mapper platforms can provide temporal and spatial global coverage beginning in the mid-1980s. Normalized Difference Tillage Index (NDTI) utilizes bands available in higher temporal coverage multispectral satellites and has proved to be robust and accurate [7,10,13]. Now that the Landsat archive exists at no cost, we can do retrospective studies. However, care must be taken when using indices, for they are affected by soil brightness, atmosphere, solar incidence angle, instrument calibration, and wavelength [14]; therefore, NDTI should be converted to residue cover (RC) to standardize across dates and locations. Supervised classifications require more data than often exists, but when available it offers adequate results [15,16].

The objectives of this study were to create remotely sensed crop residue cover maps without training data and to conduct a validation test on soil tillage intensity estimates at the field level for a test area over 10 years. All available Landsat (5, 7, and 8) imagery during the spring from 2007 to 2016 for portions of South Dakota, North Dakota, and Minnesota were screened for clouds/shadows, high soil moisture conditions, and green vegetation cover. NDTI values were converted to crop residue cover without using field-level data and then tested field-level survey data based on field polygon averages.

2. Materials and Methods

The study area along the borders of Minnesota and North Dakota and South Dakota (150,000 km^2) was selected because each state managed its EQIP program differently (Figure 1). As shown in Wallander et al. (2013) [17], South Dakota has dramatically lower participation rates with respect to tillage-related EQIP contracts as compared to North Dakota and Minnesota. We focused from the beginning of April to the end of June to capture minimum residue cover from corn and soybean fields utilizing Landsat 5 (2007–2012), Landsat 7 (2007–2016), and Landsat 8 (2014–2016). Landsat Surface Reflectance Product, on-demand version imagery is now available on the United States Geological Survey (USGS) Earth Explorer website https://earthexplorer.usgs.gov.
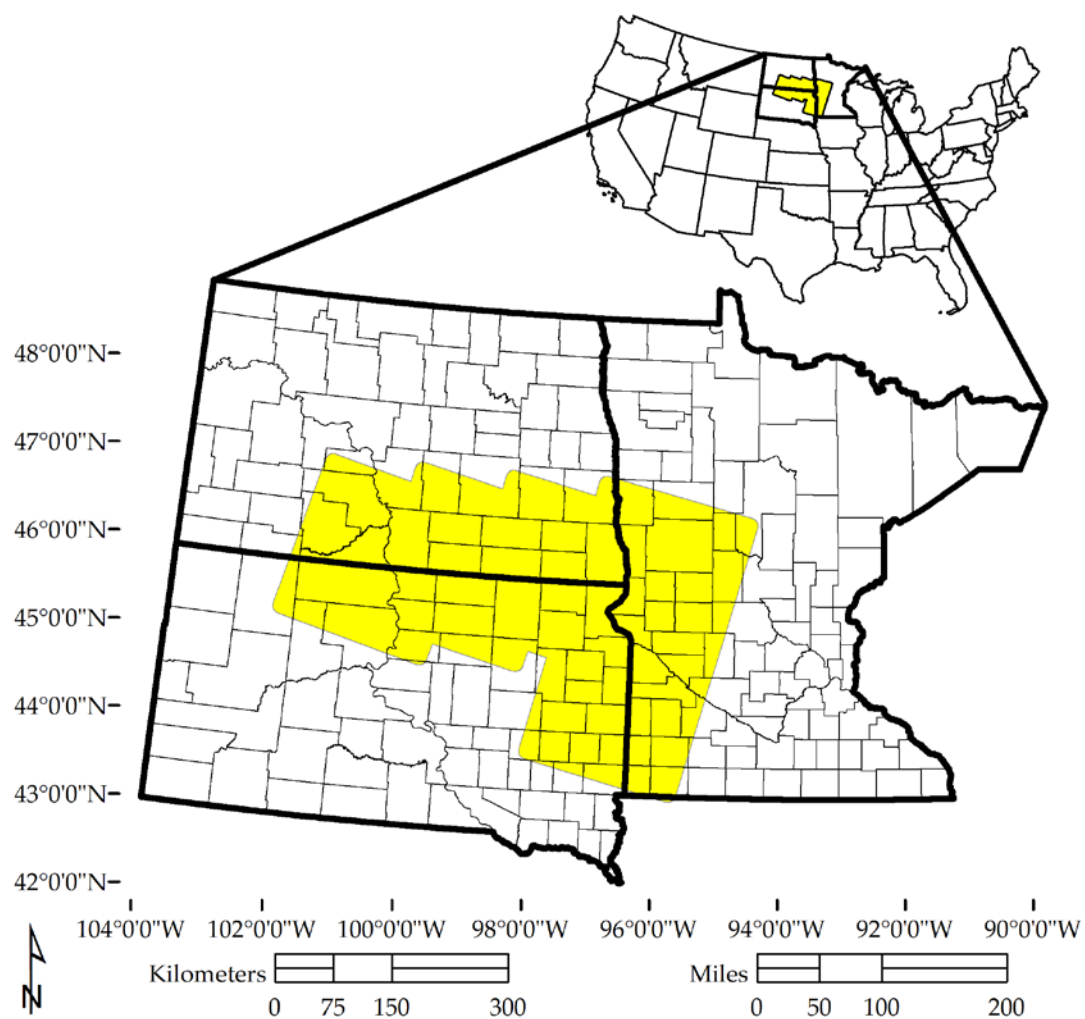
Figure 1. Location of the study area covering 150,000 km^2 in South Dakota, North Dakota, and Minnesota.

The cropland data layer (CDL) identified crop types and was used to generate polygons of persistent cropping sequences [18]. The crop residues seen in the spring were from the crops of the prior growing season. The first year the CDL was available for all three states was 2006; thus, the first year with crop residue type data was spring of 2007. The CDL can be downloaded from https://nassgeodata.gmu.edu.

High soil moisture conditions reduced the reflectance of short-wave infrared SWIR bands but increased NDTI [19]. Thus, identifying pixels of each satellite image that have high soil moisture is critical for correct estimation of crop residue using NDTI. NEXRAD Rainfall Data offers better spatial resolution (4 km) than land-based gauges and maximizes the useful pixels from each image. NEXRAD can be downloaded from the National Weather Service at http://water.weather.gov/precip/download.php. We collected NEXRAD data two days preceding the satellite overpasses to identify possible wet soils—a known complication to NDTI [20,21]. The Hydrologic Rainfall Analysis Project (HRAP) cells, a grid coordinate system used by the National Weather Service, were summed for 2-day periods prior to each satellite image. Ground-based rainfall gauges were also available, but they do not offer the spatial resolution needed to maximize the usable areas of the satellite images.

The Soil Survey Geographic Database (SSURGO) provided detailed soil survey maps at county-level scale that were used to identify zones based on soil texture that are spectrally different [22]. Data can be downloaded from https://websoilsurvey.nrcs.usda.gov. Soils were grouped into four

dominant texture classes: (1) clayey/fine/fine-silty/other/very-fine, (2) coarse-loamy/coarse-silty, (3) fine-loamy/loamy, and (4) sandy.

Digital elevation maps (DEM) were downloaded from http://ned.usgs.gov. Four slope classes were created from the 30-m DEM: 0–1, 1–2, 2–3, >3 degrees to account for variations of tillage practices on sloped land and variation of view angle that may change surface reflectance values.

2.1. Estimating Crop Residue Cover

The spectral bands for calculating Normalized Difference Vegetation Index (NDVI) [23] (Equation (1)) and Normalized Difference Tillage Index (NDTI) [7] (Equation (2)) varied slightly for each satellite (Table 1).

$$NDVI = (NIR - Red)/(NIR + Red) \tag{1}$$

$$NDTI = (SWIR1 - SWIR2)/(SWIR1 + SWIR2) \tag{2}$$

where: Red, NIR, SWIR1, and SWIR2 refer to spectral bands for Landsat 5, 7, and 8 (Table 1).

Table 1. Spectral bands of Landsat 5, 7, and 8 *.

Band	Landsat 5 TM	Landsat 7 ETM+	Landsat-8 OLI
	(nm)	(nm)	(nm)
Red	630–690	630–690	636–673
NIR	760–900	770–900	851–879
SWIR1	1550–1750	1550–1750	1566–1651
SWIR2	2080–2350	2090–2350	2107–2294

* https://landsat.usgs.gov/what-are-band-designations-landsat-satellites.

Both indices are unit-less and variable depending on atmospheric conditions, soil moisture, and/or soil color [14]. Crop residue cover (RC) was linearly related to NDTI [12,13]. We determined a relationship of RC and NDTI for each Landsat scene without the use of training data, based on the assumption that each scene has fields with low residue cover and fields with high residue cover. This differs from Zheng et al. (2013) [12], who used field data to generate a relationship between the season's minimum NDTI and measured residue cover, then used this relationship to convert minimum NDTI to RC for all scenes. We assumed residue cover varied from bare soil to a maximum residue cover (no till) as represented by the maximum and minimum NDTI values after masking, which requires no field data.

$$RC = m(NDTI) + b \tag{3}$$

$$Slope \ (m) = (RCmax - RCmin)/(NDTImax - NDTImin) \tag{4}$$

$$y\text{-intercept} \ (b) = -1* \ (slope \ (m) * NDTImin) + 0) \ \{x=0\} \tag{5}$$

Maximum and minimum NDTI (Equation (2)) values were extracted from each image and zone by using the mean plus or minus three times the standard deviation for maximum and minimum respectively to reduce noise. Maximum residue cover (RCmax) was assumed to be 85% for corn and 65% for soybean consistent with Shelton et al. (1995) [24]. Minimum residue cover (RCmin) was assumed to be 0% for both crops.

The corresponding NDTImin and NDTImax values for each image—representing 0% and 85% residue cover for corn or 65% for soybean—needed for RC calculation can be estimated with a variety of methods; here we tested five possible methods:

1. Corn and Soybean—Single value for each image regardless of crop type. Corn and soybean fields were given the same conversion and maximum residue cover was assumed to be 85%.
2. Corn or Soybean—Values for corn and soybean fields were taken separately (two equations per image) where corn has a maximum residue cover of 85% and soybean has a maximum residue cover of 65%.

3. Slope Classes—Values for the four slope classes (0–1, 1–2, 2–3, >3) resulting in four equations per image which have a mixture of corn and soybean, so the maximum residue cover was assumed to be 85%.

4. Soil Texture—Values for each soil texture (4) resulting in four equations per image which have a mixture of corn and soybean, so the maximum residue cover was assumed to be 85%.

5. Crop and Soil—Values for each crop type (2) and soil texture group (4) resulting in eight equations per image, where corn has a maximum residue cover of 85% and soybean has a maximum residue cover of 65%.

Once these residue cover estimates were calculated, the individual images were mosaicked by their minimum value to create the annual residue cover map as seen in the flow chart (Figure 2). Annual residue estimates were determined for each field based on the satellite average on a 30-m inward buffered polygon representing consistent cropping practices—crop management units. All of the above steps were scripted using ArcMap (ESRI, Redlands, California), and additional scenes or years were easily added when they became available. The quantity and quality of the satellite overpasses were important to note when extracting the results later where more overpasses indicate higher likelihood of observing the minimum crop residue cover. A quality assurance flag, indicating the number of useful dates used to calculate the minimum RC, was included to track whether the residue estimation was reliable.
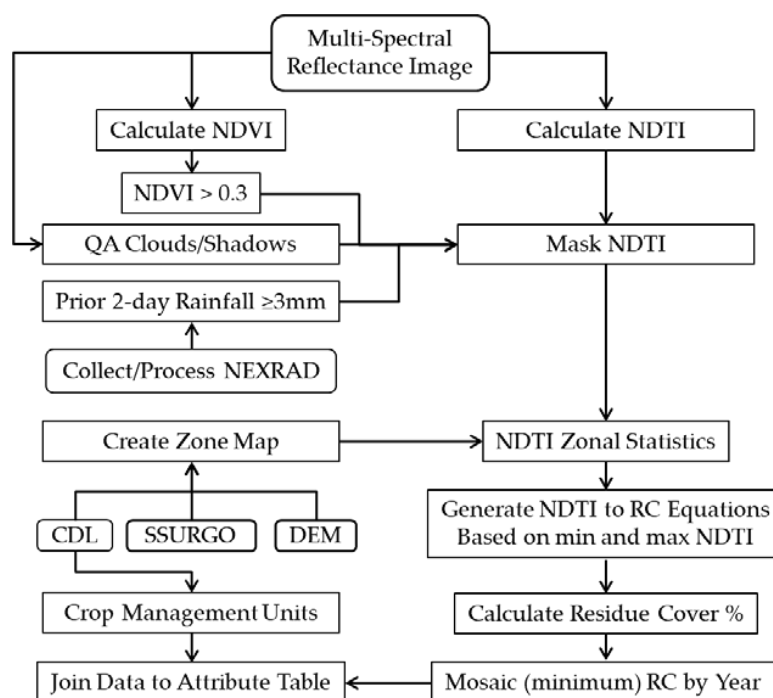

Figure 2. Flow chart for converting satellite images into percent residue cover. The resulting image values can be assigned to field polygons for further analysis.

2.2. Masking Satellite Data

Along with clouds, the usefulness of any image for NDTI is decreased by high antecedent soil moisture and green vegetation cover, i.e., crops or weeds. As soil moisture and green vegetation increase, reflectance of SWIR1 and SWIR2 decrease at slightly different rates which causes NDTI to increase [19]. We eliminated pixels that: (1) had clouds or cloud shadows, (2) had high soil moisture, and/or (3) had green vegetation over a threshold. Our goal was to maximize the usable area of each image (Figure 3). Clouds and cloud shadows were masked based on the quality assurance images provided with each scene. Most images showing 100% cloud cover were not downloaded. Then,

using the 4 km NEXRAD precipitation estimate, we masked out areas that received more than 3 mm rain during the prior 2 days as the soil moisture may have been too high. Finally, we used NDVI to mask out cropland with green vegetation. NDVI > 0.3 was selected as an average value that applies to all scenes, which was consistent with Daughtry et al. (2005) [25].
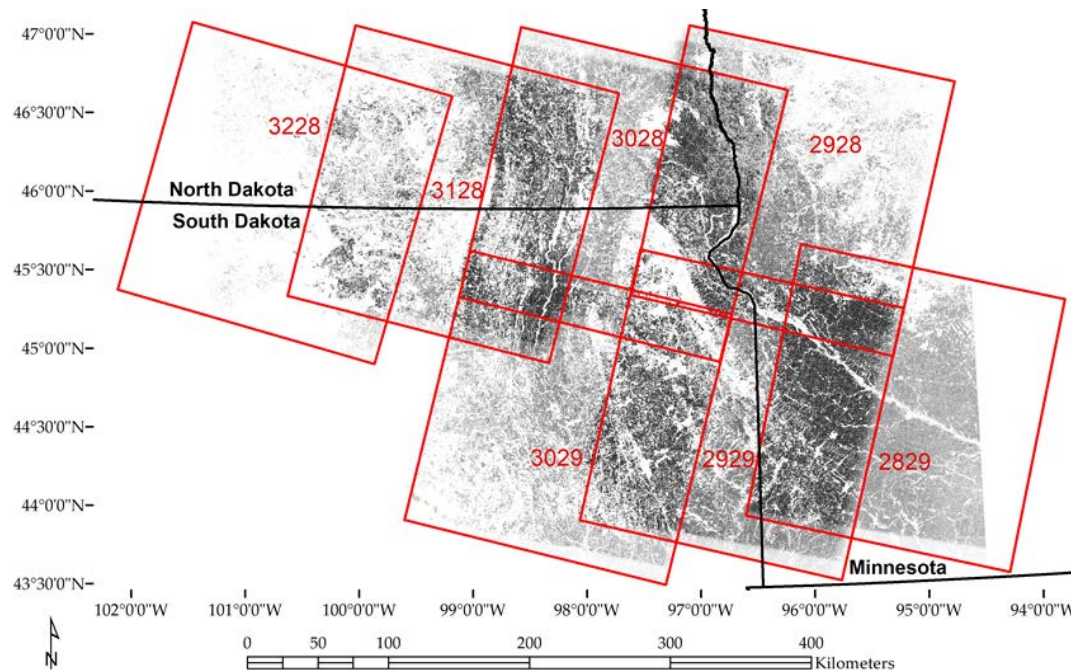

Figure 3. Locations of Landsat scene by path and row for the study area. Note east-west overlap areas. Black to white gradient indicates accumulative number of suitable images for 2007 from high counts (13) to low counts (0), respectively.

Maximum green vegetation (GV) is assumed to be 85%. By using spatially discrete masks—clouds/shadows, NEXRAD 2-day sum > 3.0 mm, and NDVI > 0.3—the largest amount of an image was maintained to increase the chances of catching the minimum NDTI while not introducing errors (Figure 4). This allowed for corn fields to use an earlier image value while keeping the rest of the image for soybean that was planted later.
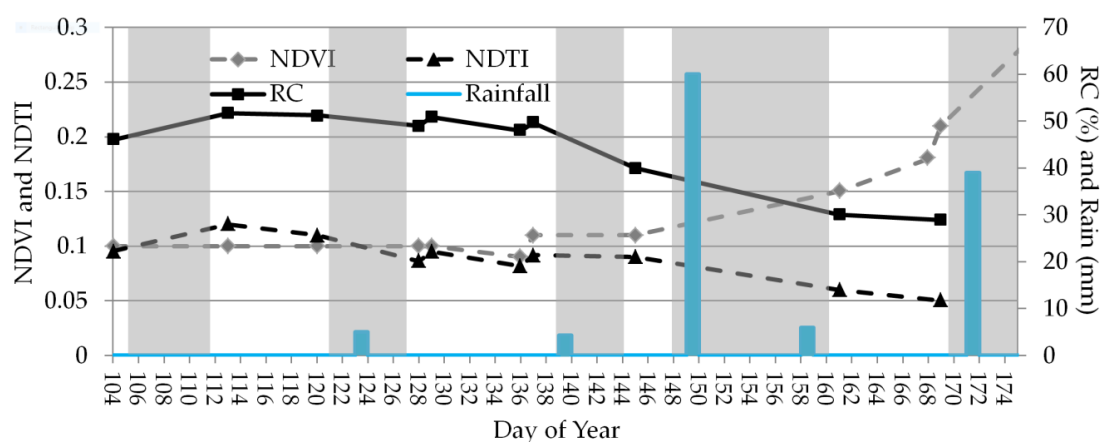

Figure 4. Synthetic representation of Normalized Difference Tillage Index (NDTI) (black dashed line) and Normalized Difference Vegetation Index (NDVI) (gray dashed line) over time and how it is masked and converted to residue cover (RC) (solid black line). The gray zones show where it is masked due to either clouds, high soil moisture (rainfall 2-day sum > 3mm, blue line), or NDVI > 0.3.

2.3. Survey Data

The validation data consisted of field-level residue estimates supplied from USDA's Agricultural Resources Management Survey (ARMS). The field-level version of ARMS collected data about management decisions for randomly selected fields growing targeted commodities (https://www.ers.usda.gov/data-products/arms-farm-financial-and-crop-production-practices/). For this analysis, we utilized the 2010 and 2016 corn and 2012 soybean versions of the survey. Detailed information on all machine operations (tillage and planting) in combination with knowledge of the prior crop were used to estimate the crop residue, as a percentage, prior to planting [24]. This method provided an estimate that captured starting residue based on crop type and subsequent residue degradation by type of machinery operation, but did not reflect other factors, such as abnormal weather or variations in machinery settings [24,26]. In addition, the survey collected self-reported use of conservation tillage categories for five years of history on each field. It did not collect the detailed information on machinery operations in those earlier years, but served as a simple two-category test of above or below 30% residue for more years. The ranges of values given for computing decreases in residue cover for tillage operations residue degradation were as much as 40% with a median of 20% [24]. Selection of the value within the given range was left up to the survey-taker and may account for some of the variation in machine settings or speed. In the six example residue calculations given in Dickey et al. (1986) [26] and Shelton et al. (1995) [24], their estimations varied as much as 31%, depending on whether the low or high value was used from the given range, and typically varied by ±25%. This variation could erroneously shift a field from being classified as conservation tillage (greater than 30% residue cover) to conventional tillage. The inaccuracies presented here may not be related to poor satellite-derived estimates but rather poor comparison data. To address this, we also report varying the ARMS data points by a conservative ±5% when it was near the conventional/conservation tillage threshold at 30% residue.

The satellite residue cover results were summarized to field polygons representing consistent crop management unit and validated ex post facto using field-level residue cover data provided by ARMS. Due to the potential for some survey observations to have incorrect geolocation, the validation exercise used alternative sets of survey observations based on the extent to which crop history and field acreage could be corroborated against other sources such as the CDL. With the survey estimates taken to be accurately located and reported estimates of tillage, the user, producer, and overall accuracies of the satellite-based tillage estimates were calculated for annual maps. Classification accuracy was assessed using the Kappa statistic, which was based on the difference between the actual agreement in the error matrix and the chance agreement [27]. For this study, a Kappa statistic ≥0.40 was desired. Data include 2010 and 2016 corn fields and 2012 soybean fields with additional information about no-till practices for five years preceding the survey data on those specific fields.

2.4. Crop Management Units

In order to compile spatial and tabular data from multiple sources, the vector polygon format provided a framework to standardize the data for further analyses. Spatial and non-spatial data at multiple scales and different time periods can be stripped of location information for safe sharing among researchers. Most spatial datasets are produced in raster format (grid-based), which are useful in disseminating a single attribute with relatively small file sizes. However, these files are limited to a predetermined spatial resolution and as a result, analyses with other datasets or formats may be difficult. For agricultural studies, there is a need to delineate polygons that describe areas of consistent cropping practices over time, so that data in various formats can be compiled.

Correct field extents establish zones of homogenous strata that are static between years, offering zones to extract data values, likely in raster format. Then values can be assigned to a polygon using statistical criteria depending on need (e.g., mean, median, majority, etc.). This will reduce imposed errors that may affect analysis by reducing spurious data points within a zone. For example, a field may contain thousands of cells from a raster data source but only one point where in situ data was collected.

Instead of selecting one value from the raster to represent that field, it is best to statistically describe the whole zone that the point data represents, thus reducing noise. Furthermore, a field could be all corn this year but might have been half corn and half soybean last year. Only when the field extents are known can descriptive statistics be computed accurately. More years that need to be represented by one zone will cause the number of polygons to increase as producers choose different operations for one side of a field than the other. Here we needed polygon delineations that represent unique static zones of cropping and operation sequences—crop management units. This may vary within a particular farm across years even though the producer remains the same.

Currently, the common land unit (CLU) product by the USDA Natural Resources Conservation Service (NRCS) is a polygon-based representation of agricultural fields that were digitized from ostensibly permanent land features such as roads, waterways, and/or tree and fence lines from aerial photos. For a particular year, it is the smallest area that has a discrete boundary and consistent land cover. There are drawbacks to the CLUs that could make it the wrong choice for describing fields: (1) CLUs are inaccessible to researchers outside USDA due to privacy issues; (2) CLUs do not capture variations within a single field across years; (3) CLUs have different versions by year and sometimes have incomplete coverage; (4) CLUs may contain errors such as repeat or missing polygons, misaligned boundaries). Often researchers that have access to these files spend time dividing polygons to represent other years to meet their needs. This can be very time-intensive and is usually done manually. To overcome these issues, it may be beneficial to generate one's own polygons that match the years needed within the study. These polygons can be scaled to meet individual needs of the researcher while representing zones of unique crop sequences and capturing variations within fields, independent of owner/operator for as many years as the Cropland Data Layer exists (now often over 10 years for some locations in the U.S.). Perhaps the best reason to delineate one's own polygons is that abides by the Food, Conservation, and Energy Act of 2008, Title I—Commodity Programs, Subtitle F—Administration, Section 1619.

Other studies have focused on describing unique cropping practices [28,29]. These methods were effective in distilling dominant cropping sequences from 3 to 6 years of CDL down to 1- to 3-year rotations to describe the land in as few rotations as needed, but do not describe sequences on a field-level basis. Here, 11 years of the CDL were used to create the distinct zones specific to that area and aggregated to delineate the fields. The steps were to filter and simplify codes on each CDL year to remove as much noise as possible then stack desired years and re-impose road and rail network. These features were lost in filtering and indicated boundaries that divide fields that should be seen in the final result. The resulting polygons were then dilated and eroded to remove islands within fields. This resulted in a final polygon that was inward buffered 30 m from the full extent so that descriptive statistics did not include field edges—often treated differently than the rest of the field due to increased vehicle traffic for field operations. The polygons served as zones to extract values from other datasets based on desired descriptive statistics. The resulting crop management units are polygons that represent areas of similar cropping sequences for the designated years, frequently dividing a single CLU into many polygons.

3. Results

Timing and availability of the satellite images are always the weak point of a remote sensing study. Some locations had as many as seven useful images per year, while other areas had only one useful image. The east-west overlap area created by adjacent Landsat paths (Figure 3) increased number of useful images to as high as 13 which increased the opportunities for capturing the minimum crop residue cover accurately. Conversely, one of these overlapped areas only had two useable images and likely missed capturing the minimum residue feature. This was uncommon and most areas had more useable images (Table 2), which increased the reliability of the results.

Table 2. Summary of satellite images and the accumulation in the east-west overlap occurrences and summary of image acquisition dates (day of year). Dark gray fill indicates where the image is more than 100% to 75% masked, Light gray fill indicates where the image is 75% to 50% masked, and non-filled cells are greater than 50% remaining for NDTI.

Path Row	32 28	Overlap Count	31 28	Overlap Count	30 28	Overlap Count	29 28	30 29	Overlap Count	29 29	Overlap Count	28 29	Total
2007	2	5	3	7	4	8	4	3	9	6	13	7	29
DOY	118 174		143 159 175		128 136 160 176		129 137 161 177	160 168 176		97 121 129 137 145 161		98 122 130 138 162 170 178	
2008	2	5	3	5	2	5	3	0	4	4	8	4	18
DOY	113 153		138 146 178		139 171		124 132 164			100 124 132 164		125 141 173 181	
2009	4	7	3	4	1	3	2	5	6	1	4	3	19
DOY	123 139 147 155		148 164 180		181		142 150	125 133 149 173 181		150		127 151 175	
2010	3	8	5	8	3	7	4	3	8	5	9	4	27
DOY	110 134 150		111 143 159 167 175		104 112 128		97 105 137	104 112 128		97 105 137 153 169		98 106 138 170	
2011	4	9	5	7	2	4	2	3	5	2	3	1	19
DOY	97 113 121 145		114 122 146 154 170		115 123		156 180	123 131 179		156 180		101	

Table 2. *Cont.*

Path Row	32 28	Overlap Count	31 28	Overlap Count	30 28	Overlap Count	29 28	30 29	Overlap Count	29 29	Overlap Count	28 29	Total
2012	3	5	2	4	2	4	2	2	3	1	2	1	13
DOY	92 116 156		101 133		94 174		135 167	94 174		135		176	
2013						(insufficient satellite coverage)							
2014	1	3	2	5	3	8	5	2	6	4	6	2	19
DOY	121		98 114		107 163 171		100 140 148 172 180	163 171		100 148 172 180		125 157	
2015	4	9	5	9	4	7	3	4	7	3	6	3	26
DOY	100 140 172 180		93 101 117 141 165		142 150 158 166		103 119 159	94 118 142 150		103 119 159		112 160 176	
2016	3	10	7	13	6	9	3	6	10	4	6	2	31
DOY	127 159 167		104 112 120 128 136 144 168		113 129 137 145 161 169		122 138 154	105 113 129 137 145 161		122 138 154 170		139 147	
Total	26		35		27		28	28		30		27	201

After masking, some images were completely eliminated (Figure 5). The images were masked based on any of the following occurrences: clouds/shadows, high soil moisture (rainfall accumulation > 3 mm for two days prior), and/or growing vegetation (NDVI > 0.3).
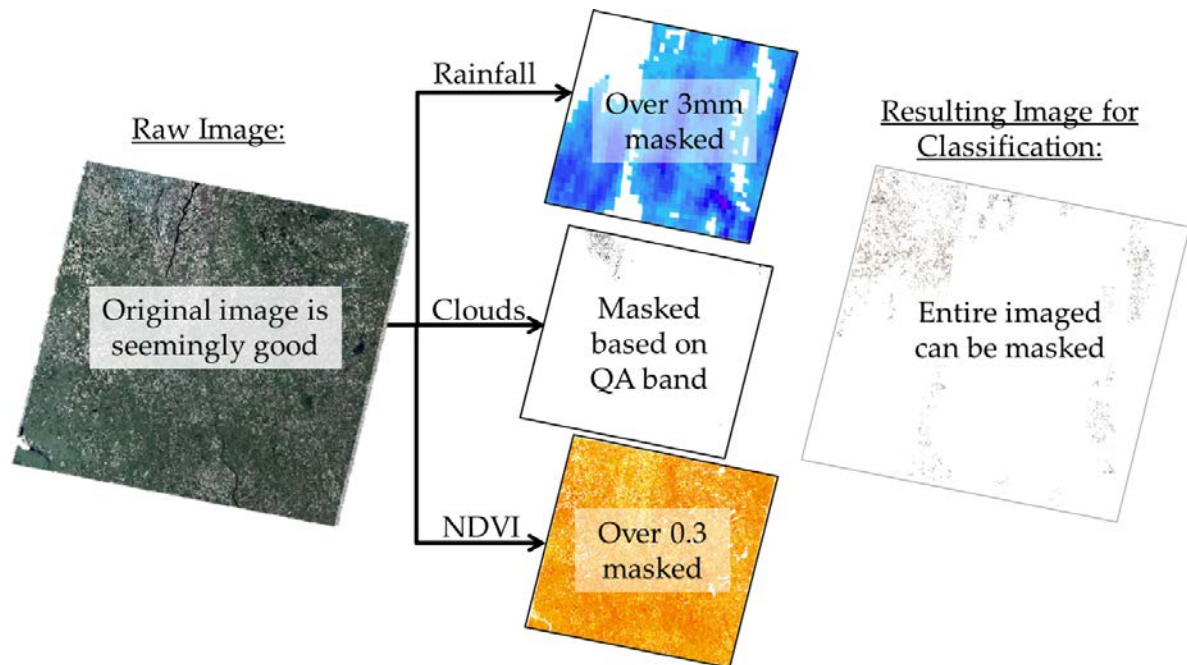

Figure 5. The result of masking one satellite image from 2007 based high soil moisture (2-day rainfall sum over 3 mm), clouds or cloud shadows, and/or 3) green vegetation (NDVI over 0.3). (Conceptual figure and not to scale, a Landsat image size is 185 km × 180 km.).

The proportion of the remaining image useful for the next step is represented in Table 2. Typically, as the dates extend into June, less of each image was useable, which was likely attributed to crops growing and more of the image masked by NDVI. However, the chances of using an NDTI value before planting increased because some of the most complete/useful images were from earlier in the spring and were not necessarily the lowest residue cover.

The resulting images were processed to determine the conversion (Equations (4) and (5)) (slope and y-intercept) from the NDTI value to residue cover percentage based on the five methods: (1) Single value for each scene where corn and soybean fields were given the same conversion; (2) Values for corn and soybean fields were taken separately (two equations per image); (3) Values for each soil texture class resulting in four equations per image; (4) Values for the four field slope classes (0–1, 1–2, 2–3, >3) (four equations per image); (5) Values for each crop type and soil texture (eight equations per image). The values for slope and y-intercept from (Equations (4) and (5)) varied within each method and across the methods (Figure 6). Slope values by method averaged 350 to 415, and y-intercept values averaged 7.8 to 10.9. Indices were affected by soil brightness, atmosphere, solar incidence angle, instrument calibration, and wavelength [15]; so, by taking each image separately and determining a conversion to residue cover percentage, these factors were reduced. Zheng et al., (2012) [30] determined a slope of 755 and y-intercept of 5.4 from field-level data and then applied that relationship to all minimum NDTI (minNDTI) values across many Landsat images for their study area in Indiana. Due to the scene to scene variation observed, we contend that our method is more robust across satellite platforms, atmospheric conditions, and requires no field data to determine relationship. Landsat 8 had consistently higher slope values than the other satellites (+52 and +68 on average for Landsat 5 and 7, respectively), and had consistently lower y-intercepts on average (−3.2 and −2.1 for Landsat 5 and 7, respectively).
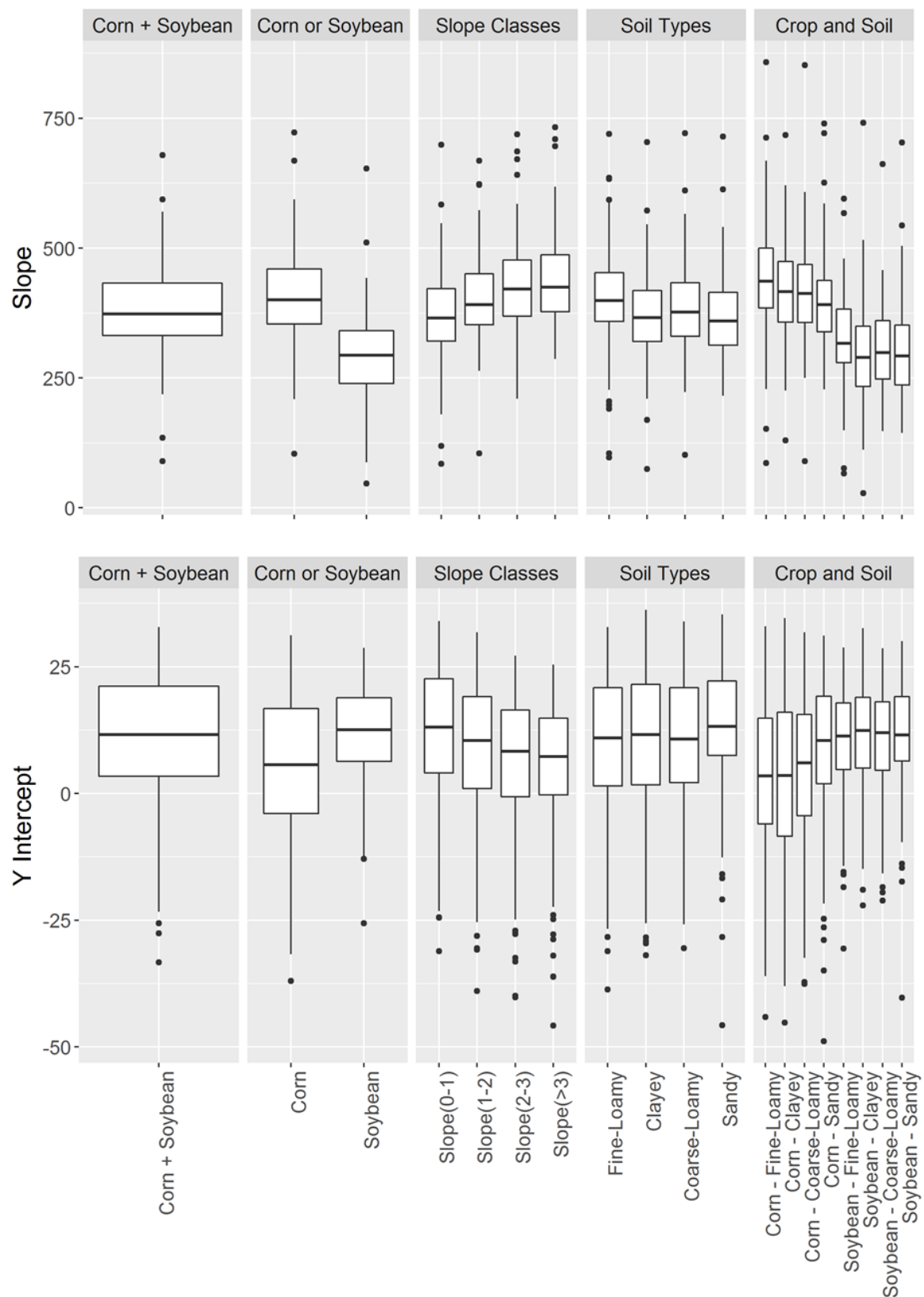
Figure 6. Ranges of slope (**top**) and Y-intercept (**bottom**) used to convert NDTI to percent residue cover for the five methods. Top and bottom whiskers denote 90th and 10th percentiles, respectively. Top and bottom of the boxes denote the 75th and 25th percentiles, respectively. Lines within boxes denote median values. Black dots denote outliers below and above the 10th and 90th percentiles, respectively.

Fields included in the ARMS data were identified in the classified Landsat images and verified with acreage and cropping sequences to ensure correct geolocation of the survey point. This provided 200 fields for comparison to the Landsat-based residue cover maps; 65 and 72 corn fields in 2010 and 2016, respectively, 63 soybean fields in 2012, and 219 additional fields that reported no-till operations for the past five years. From the crop management units, crop residue cover was assessed for more than 500,000 corn and soybean fields using the Landsat data in the study area. User-producer confusion matrices were created for each of the datasets; 2010 corn, 2012 soybean, 2016 corn, and overall accuracy is reported (Table 3). The residue cover calculated by the satellites had between one and 13 overpasses.
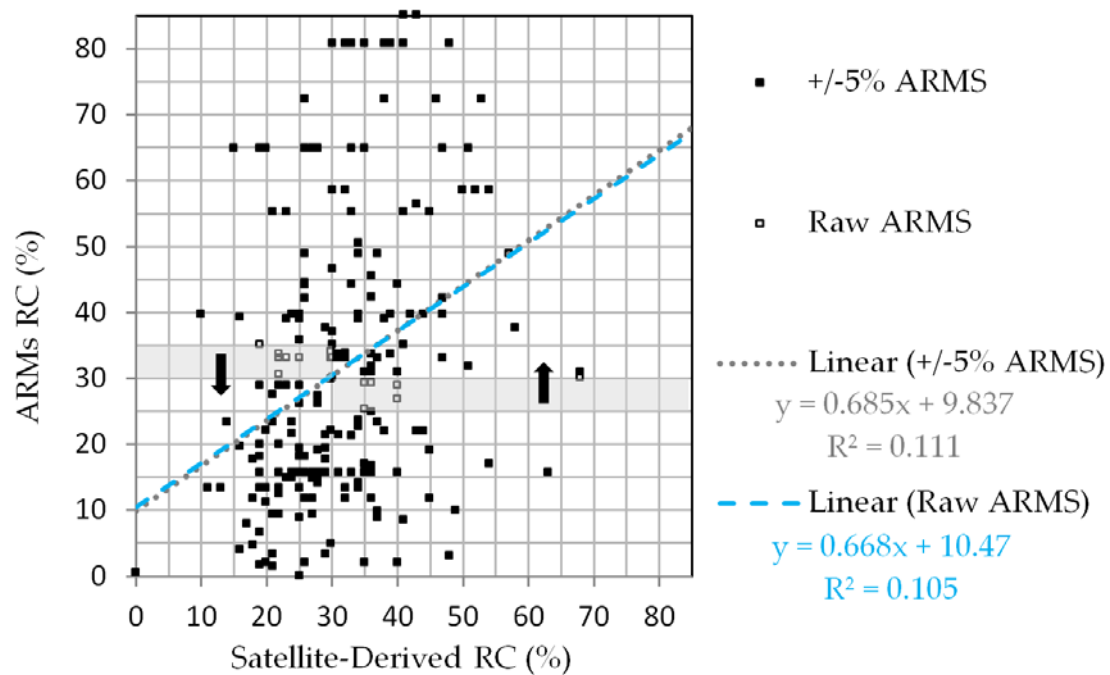
Table 3. Summary statistics for five methods of assessing crop residue cover for all ARMS fields (top), for fields with quality assurance flag >3 Landsat images (middle), and for fields with quality assurance flag ≥4 Landsat images (bottom). Both raw ARMS data and ARMS ±5% at the category boundary are included.

	1	2	3	4	5
	Corn and Soybean	Corn or Soybean	Slope Classes	Soil Texture	Crop and Soil Texture
	All data points n = 200 fields				
Accuracy (Raw)	61%	67%	62%	60%	65%
Kappa (Raw)	0.24	0.33	0.25	0.22	0.29
Accuracy (ARMs ±5%)	68%	73%	68%	68%	71%
Kappa (ARMs ±5%)	0.36	0.45	0.37	0.37	0.42
R^2	0.12	0.11	0.12	0.12	0.09
Slope	0.70	0.69	0.68	0.70	0.60
Y-intercept	6.48	9.84	7.24	6.60	12.61
	Quality Assurance ≥ 3 n = 121 fields				
Accuracy (Raw)	62%	70%	61%	60%	69%
Kappa (Raw)	0.25	0.38	0.25	0.22	0.34
Accuracy (ARMs ±5%)	68%	74%	67%	67%	75%
Kappa (ARMs ±5%)	0.37	0.42	0.36	0.36	0.46
R^2	0.19	0.18	0.17	0.17	0.16
Slope	0.93	0.98	0.90	0.93	0.89
Y-int	−0.35	1.95	0.56	−0.47	4.51
	Quality Assurance ≥ 4 n = 72 fields				
Accuracy (Raw)	72%	78%	69%	72%	72%
Kappa (Raw)	0.43	0.54	0.39	0.44	0.42
Accuracy (ARMs ±5%)	81%	85%	78%	81%	82%
Kappa (ARMs ±5%)	0.60	0.66	0.55	0.60	0.59
R^2	0.28	0.25	0.26	0.26	0.19
Slope	1.16	1.24	1.14	1.22	1.01
Y-int	−3.60	−1.05	−3.22	−5.44	4.75

This also addressed the issue from converting continuous data to discrete categories (conventional/conservation tillage) at the 30% boundary. For example, a reported residue value of 29% just missed the conservation tillage category and skewed the results artificially. Both results, raw and ±5% adjusted, were included in the results. The satellite-derived residue estimates were not adjusted by ±5%.

In general, the overall accuracies for the raw ARMS results for the five methods varied from 61% to 67% with all Kappa statistics under 0.40, indicating fair to poor agreement (Table 3). These results were improved when the ARMS data were allowed a ±5% buffer at the 30% threshold to 68% to 73% with Kappa statistics over 0.40 for two of the methods. The R^2 and slope were minimally affected by adjusting the ARMS data by ±5% (Figure 7). Including a quality assurance flag, the user can inspect low temporal overpass fields as probable error. Using this quality assurance flag, only fields with greater than three and four satellite overpasses were used, which eliminated spurious data due to poor satellite coverage or timing and improved all statistics. Slope was closest to one for quality assurance greater than three, but the R^2 were best for quality assurance greater than four. All methods had Kappa statistics >0.40 when the quality assurance value was above four and when the ARMS data had ±5% residue at the category boundary.
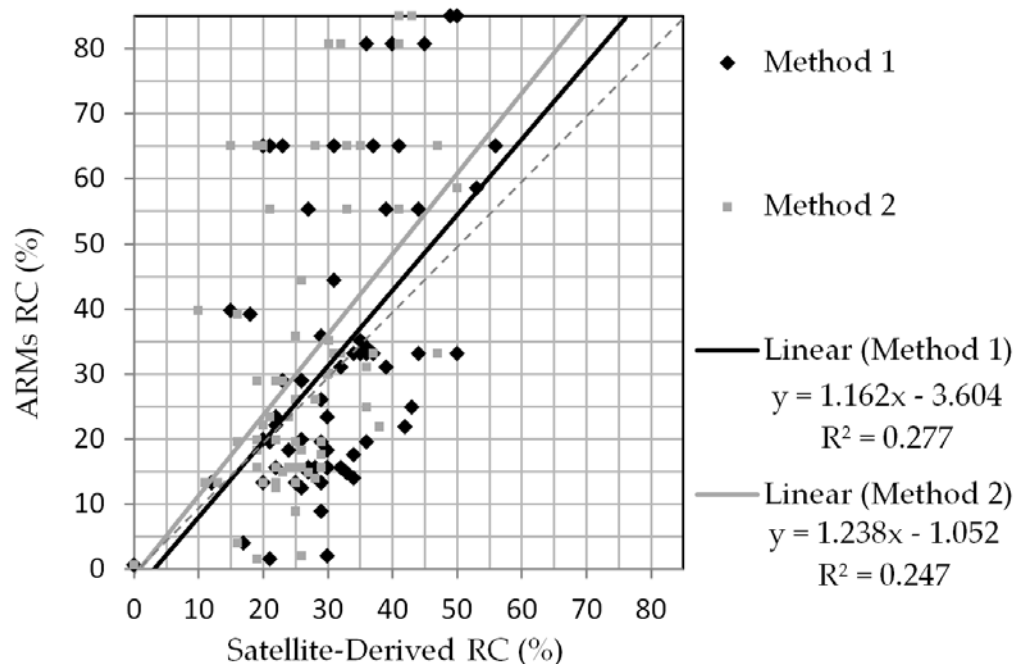
A.

B.

Figure 7. (**A**). For Method 2 adjusting the ARMS data near the category boundary by ±5% did not affect R^2 (gray dotted and blue dashed lines overlaying each other). (**B**). Using points four or more satellite overpasses for Methods 1 and 2 increases R^2 and slope is closer to one-to-one (black line).

The best performing method was Method 2, where corn and soybean each receive their own scaling equations; accuracy and Kappa statistics increased from 67%, 0.33 to 85%, 0.66. Method 5 was the next best performing method, crop, and soil, but because there were eight equations for each image, and each image could be drastically reduced by masking, the chances of capturing accurate

low and high residue field for scaling NDTI to residue cover was low. This method would be best for larger footprint satellite platforms and might require mostly unmasked images for best results. And given the slope was closest to one, this method should be used for large areas, but would result in many more soil classes and would require more effort for possibly minimal improvement. The worst performing method was Method 3, where four slope classes were used. This was because, although there was variation in the distribution of lower tilled fields on steeper slopes, there were often few fields with very low residue in steep classes needed to generate the scaling equations, and accounting for variation in satellite view-angles, these were not as sensitive as crop type. Method 4, using soil texture classes, also proved less accurate because the size of Landsat images was not big enough that changes in soil texture influenced scaling to residue values as much as separating crop residue types.

Kappa statistic is based on the difference between the actual agreement in the error matrix and the chance agreement [25].

The no till occurrences reported in the ARMS data were expected to perform the best as estimated by the satellites, but there was limited success with the comparisons (Table 4). When the ARMS data reported no till for prior 5 years, the expected residue cover should be greater >30%. We compared satellite-based estimates to fields identified in the ARMS data, as well as the residue cover calculated using tillage operations [24] reported in the ARMS data. The ARMS data, when comparing reported residue values to reported no till operation flag, was only correct 61% of the time or 65% when ±5% was allowed near the category boundary. The satellite derived estimates were not adjusted by ±5%. Methods 1 (corn and soybean), 3 (slope class), and 4 (soil texture) were best at 76% correct capturing on-till occurrences reported by ARMs. However, when all the data points were used, the results were better than when the quality assurance flag was greater than four by about 4%. This could be related to poor satellite estimations, but it is likely due to the potentially omitted variables, noted above in the residue estimates built from the ARMS data. However, methods that produce overall higher residue values would perform better due to omission of data indicating conventionally tilled fields.

Table 4. Summary statistics of accuracy for the five remote sensing methods and ARMS estimated residue cover compared to ARMS data reporting no-till operations (testing if greater than 30% residue).

	Number of ARM Fields Reported No-Till	1 Corn + Soybean	2 Corn or Soybean	3 Slope Class	4 Soil Texture	5 Crop and Soil Texture	Number of ARM Fields Over 30% and No-Till	ARMS Residue Estimate	ARMS Residue +5%
2007	10	90%	70%	80%	90%	70%			
2008	28	75%	68%	75%	75%	68%			
2009	28	82%	79%	82%	82%	75%			
2010	36	72%	61%	69%	69%	56%	15	80%	80%
2011	21	76%	52%	76%	76%	57%			
2012	29	83%	72%	86%	83%	69%	23	78%	78%
2014	9	67%	44%	67%	67%	56%			
2015	19	84%	63%	84%	79%	63%			
2016	39	67%	49%	69%	69%	49%	39	44%	48%
All Data	219	76%	63%	76%	76%	62%	77	61%	65%
QA ≥ 4 *	100	72%	58%	72%	73%	58%	18	70%	72%

* QA (quality assurance) is the number of useful dates used to calculate the minimum RC (higher is better).

4. Discussion

While the discrete residue classification described above had classification errors larger than many studies built on training data, we also looked at how the satellite-based estimates of residue, absent a discrete tillage determination, compared to the survey-based estimates of residue in 2010, 2012, and 2016. Given the dearth of historical validation data, if the satellite data had any statistically significant ability to predict residue levels, then satellite-based estimates contained some meaningful signal of actual residue. For many statistical studies of program effects, this implies that statistical controls could be used to measure what signal was available. This could be an extremely valuable contribution to historical studies because satellite data could be compiled for hundreds of thousands of fields in the study area, whereas historical survey data is limited to a few hundred fields. The slopes

of the linear fit shown above (Table 3) were statistically significant, showing that there was meaningful signal in the satellite-based estimates.

Comparisons to the ARMS data were mixed. Caution should be exercised when using this type of data, and it should be regarded as a rough guide at best. There were several potential sources of error in the validation data: (1) inaccurate geolocation of fields; (2) self-reporting errors in tillage operations; and (3) crop residue cover values derived from a method known to create a large range to its residue cover estimates. The issue of poor geolocation was handled by eliminating fields that did not align with acreage and past cropping sequences observed at field level using CDL data [18]. The issue of the ARMS data using a method that estimates residue cover values from field operations which produces a wide range should be a concern. However, the number of fields that now have residue cover estimates is much greater than the ARMS data alone, and when used in the econometric models within USDA ERS, the larger sample size outweighs the noise. While this approach may have noisier estimates than other methods that have been limited to small regions with training data, this approach will allow analyses of historical data and data covering large regions. For program evaluation, the extremely large gain in signal is likely to outweigh any increase in noise.

To better assess the crop residue cover at planting, more satellite images are needed. For this study only two Landsats (5 and 7 or 7 and 8) were available in a single year. Note that 2013 was not computed due to lack of operating satellites. There were only nine total images for the study area in 2013 before any masking, so residue cover was not computed because it would likely not capture the minimum residue value accurately. The locations that relied on just one satellite image, might be too high in general because it probably captured the field before all tillage and planting operations were completed, and residue cover is reduced with each operation—never increased. In 2015 and 2017 the European Space Agency launched a Sentinel-2 satellite with appropriate multispectral bands for crop residue monitoring and with a revisit time of 5 days for the two satellites. Harmonized Landsat-8 and Sentinel-2 surface reflectance data will provide new opportunities for data every 3–5 days depending on atmospheric conditions [31] and will be widely available such as the Landsat archive [32]. This would increase the number of fields that have more than four useable images and should result in better maps of crop residue cover and soil tillage intensity in the future.

The assumption that high and low residue fields exist within an image decreases (1) when the number of equations needed to estimate residue cover increases; or (2) when the equation classes do not align to high and low residue fields; or (3) when too much of an image is masked out. For example, in method 3, slope classes, typically steep fields would be less tilled for conservation reasons and finding a low residue field to supply the equation might be limited. Additionally, in method 4, soil texture, highly erodible soils would be tilled less for conservation reasons resulting in the same issue. Method 5, crop and soil texture, suffers from two issues where low residue might be absent in erodible soils and the number of equations limit the chances of finding high and low fields. All of these examples can be made more sensitive when the images are masked due to clouds, soil moisture, and/or green cover. Greater care can be taken to limit the use of these classes and images when these situations arise.

5. Conclusions

In this study, we estimated historical crop residue cover with multi-spectral and multi-temporal satellite images using Normalized Difference Tillage Index (NDTI) without the use of training data. NDTI can be affected by (1) clouds or cloud shadows, (2) high soil moisture, and (3) green vegetation; but by masking areas of each image influenced by these factors, we maximized the useable area of each image while not including erroneous NDTI results. Converting the NDTI values to residue cover was dependent on crop residue type, soil type, and landscape slope, but crop residue type yielded the best results. We tested the influence of the other factors and found residue cover type was the most important to address. It had 85% accuracy when fields with greater than four overpasses were used and when tolerance was given to the strict tillage category boundary created by ARMS data with known range in their residue values. However, when using satellite platforms with larger image

footprints, including soil type would be beneficial, but the improvement may not be significant enough to justify the extra effort.

References

1. CTIC (Conservation Technology Information Center). *Procedures for Using the Cropland Roadside Transect Survey for Obtaining Tillage Crop Residue Data*; Conservation Technology Information Center, Purdue University: West Lafayette, IN, USA, 2009; Available online: http://www.ctic.org (accessed on 17 August 2020).

2. Lal, R.; Sobeski, T.M.; Iivary, T.; Kimble, J.M. *Soil Degradation in the United States*; Lewis Publishers: Boca Raton, FL, USA, 2004.

3. Daughtry, C.S.T.; Doraiswamy, P.C.; Hunt, E.R.; Stern, A.J.; McMurtrey, J.E.; Prueger, J.H. Remote sensing of crop residue cover and soil tillage intensity. *Soil Tillage Res.* **2006**, *91*, 101–108.

4. Gowda, P.H.; Dalzell, B.J.; Mulla, D.J.; Kollman, F. Mapping tillage practices with Landsat thematic mapper based logistic regression models. *J. Soil Water Conserv.* **2001**, *56*, 91–96.

5. Sullivan, D.G.; Strickland, T.C.; Masters, M.H. Satellite mapping of conservation tillage adoption in the Little River experimental watershed, Georgia. *J. Soil Water Conserv.* **2008**, *63*, 112–119. [CrossRef]

6. Thoma, D.P.; Gupta, S.C.; Bauer, M.E. Evaluation of optical remote sensing models for crop residue cover assessment. *J. Soil Water Conserv.* **2004**, *59*, 224–233.

7. Van Deventer, P.; Ward, D.; Gowda, P.H.M.; Lyon, J.G. Using thematic mapper data to identify contrasting soil plains and tillage practices. *Photogramm. Eng. Remote Sens.* **1997**, *63*, 87–93.

8. Daughtry, C.S.T. Agroclimatology: Discriminating crop residues from soil by shortwave infrared reflectance. *Agron. J.* **2001**, *93*, 125–131. [CrossRef]

9. Bannari, A.; Pacheco, A.; Staenz, K.; McNairn, H.; Omari, K. Estimating and mapping crop residues cover on agricultural lands using hyperspectral and IKONOS data. *Remote Sens. Environ.* **2006**, *104*, 447–459. [CrossRef]

10. Serbin, G.; Hunt, E.R.; Daughtry, C.S.T.; McCarty, G.W.; Doraiswamy, P.C. An improved ASTER index for remote sensing of crop residue. *Remote Sens.* **2009**, *1*, 971–991.

11. Hively, W.D.; Lamb, B.T.; Daughtry, C.S.T.; Shermeyer, J.; McCarty, G.W.; Quemada, M. Mapping Crop Residue and Tillage Intensity Using WorldView-3 Satellite Shortwave Infrared Residue Indices. *Remote Sens.* **2018**, *10*, 1657. [CrossRef]

12. Zheng, B.; Campbell, J.B.; Serbin, G.; Daughtry, C.S.T. Multitemporal remote sensing of crop residue cover and tillage practices: A validation of the minNDTI strategy in the United States. *J. Soil Water Conserv.* **2013**, *68*, 120–131. [CrossRef]

13. Zheng, B.; Campbell, J.B.; Serbin, G.; Galbraith, J.M. Remote sensing of crop residue and tillage practices: Present capabilities and future prospects. *Soil Tillage Res.* **2014**, *138*, 26–34. [CrossRef]

14. Jenson, J.R. *Remote Sensing of the Environment: An Earth Resource Perspective*; Pearson Prentice Hall: Upper Saddle River, NJ, USA, 2000; 592p.

15. Pacheco, A.; McNairn, H.; Smith, A.M. Multispectral indices and advanced classification techniques to detect percent residue cover over agricultural crops using Landsat data. *Proc. SPIE* **2006**, 6298. [CrossRef]

16. Beeson, P.C.; Daughtry, C.S.T.; Hunt, E.R.; Akhmedov, B.; Sadeghi, A.M.; Karlen, D.L.; Tomer, M.D. Multispectral satellite mapping of crop residue cover and tillage intensity in Iowa. *J. Soil Water Conserv.* **2016**, *71*, 385–395. [CrossRef]

17. Wallander, S.; Aillery, M.; Hellerstein, D.; Hand, M. *The Role of Conservation Programs in Drought Risk Adaptation*; U.S. Department of Agriculture, Economic Research Service: Washington, DC, USA, 2013; Report Number 148.

18. Johnson, D.M.; Mueller, R. The 2009 cropland data layer. *Photogramm. Eng. Remote Sens.* **2010**, *76*, 1201–1205.

19. Quemada, M.; Daughtry, C.S.T. Spectral indices to improve crop residue cover estimation under varying moisture conditions. *Remote Sens.* **2016**, *8*, 660. [CrossRef]

20. Daughtry, C.S.T.; Hunt, E.R., Jr. Mitigating the effects of soil and residue water contents on remotely sensed estimates of crop residue cover. *Remote Sens. Environ.* **2008**, *112*, 1647–1657. [CrossRef]

21. Quemada, M.; Hively, W.D.; Daughtry, C.S.T.; Lamb, B.T.; Shermeyer, J. Improved crop residue cover estimates obtained by coupling spectral indices for residue and moisture. *Remote Sens. Environ.* **2018**, *206*, 33–44. [CrossRef]

22. Soil Survey Staff; Natural Resources Conservation Service; United States Department of Agriculture. Web Soil Survey. Available online: https://websoilsurvey.nrcs.usda.gov/ (accessed on 29 June 2020).

23. Tucker, C.J. Red and photographic infrared linear combinations monitoring vegetation. *J. Remote Sens. Environ.* **1979**, *8*, 127–150. [CrossRef]

24. Shelton, D.P.; Smith, J.A.; Jasa, P.J.; Kanable, R. G95-1135 Estimating Percent Residue Cover Using the Calculation Method. Historical Materials from University of Nebraska-Lincoln Extension. 1995. Article Number 780. Available online: http://digitalcommons.unl.edu/extensionhist/780 (accessed on 17 August 2020).

25. Daughtry, C.S.T.; Hunt, E.R., Jr.; Doraiswamy, P.C.; McMurtrey, J.E., III. Remote sensing the spatial distribution of crop residues. *Agron. J.* **2005**, *97*, 864–871. [CrossRef]

26. Dickey, E.C.; Jasa, P.J.; Shelton, D.P. Estimating Residue Cover. Biological Systems Engineering: Papers and Publications. 1986. Article Number 255. Available online: http://digitalcommons.unl.edu/biosysengfacpub/255 (accessed on 17 August 2020).

27. Congalton, R.G.; Green, K. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, 2nd ed.; Taylor & Francis Group: Boca Raton, FL, USA, 2009.

28. Boryan, C.; Yang, Z.; Willis, P.; Di, L. Developing Crop Specific Area Frame Stratifications based on Geospatial Crop Frequency and Cultivation Data Layers. *J. Integr. Agric.* **2017**, *16*, 312–323. [CrossRef]

29. Sahajpal, R.; Zhang, X.; Izaurralde, R.C.; Gelfand, I.; Hurtt, G.C. Identifying representative crop rotation patterns and grassland loss in the US Western Corn Belt. *Comput. Electron. Agric.* **2014**, *2014. 108*, 173–182. [CrossRef]

30. Zheng, B.; Campbell, J.B.; de Beurs, K.M. Remote sensing of crop residue cover using multi-temporal Landsat imagery. *Remote Sens. Environ.* **2012**, *117*, 177–183. [CrossRef]

31. Claverie, M.; Ju, J.; Masek, J.G.; Dungan, J.L.; Vermote, E.F.; Roger, J.-C.; Skakun, S.V.; Justice, C. The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote. Sens. Environ.* **2018**, *219*, 145–161. [CrossRef]

32. Zhou, Q.; Rover, J.; Brown, J.; Worstell, B.; Howard, D.; Wu, Z.; Gallant, A.L.; Rundquist, B.; Burke, M. Monitoring Landscape Dynamics in Central U.S. Grasslands with Harmonized Landsat-8 and Sentinel-2 Time Series Data. *Remote Sens.* **2019**, *11*, 328. [CrossRef]