*Article*

# Estimation of Crop Residue Cover Utilizing Multiple Ground Truth Survey Techniques and Multi-Satellite Regression Models

**Forrest Williams [1], Brian Gelder [1] , DeAnn Presley [2] , Bryce Pape [3],* and Andrea Einck [1]**

[1] Department of Agricultural and Biosystems Engineering, Iowa State University, Ames, IA 50011, USA; forrestw@iastate.edu (F.W.); bkgelder@iastate.edu (B.G.); ameinck@gmail.com (A.E.)
[2] Department of Agronomy, Kansas State University, Manhattan, KS 66506, USA; deann@ksu.edu
[3] Department of Community and Regional Planning, Iowa State University, Ames, IA 50011, USA
* Correspondence: bppape@iastate.edu

**Abstract:** Soil erosion within agricultural landscapes has significant environmental and economic impacts and is strongly driven by reduced residue cover in agricultural fields. Large-area soil erosion models such as the Daily Erosion Project are important tools for understanding the patterns of soil erosion, but they rely on the accurate estimation of crop residue cover over large regions to infer the tillage practices, an erosion model input. Remote sensing analyses are becoming accepted as a reliable way to estimate crop residue cover, but most use localized training datasets that may not scale well outside small study areas. An alternative source of training data may be commonly conducted tillage surveys that capture information via rapid "windshield" surveys. In this study, we utilized the Google Earth Engine to assess the utility of three crop residue survey types (windshield tillage surveys, windshield binned residue surveys, and photo analysis surveys) and one synthetic survey (retroactively binned photo analysis data) as sources of training data for crop residue cover regressions. We found that neither windshield-based survey method was able to produce reliable regressions but that they can produce reasonable distinctions between low-residue and high-residue fields. On the other hand, both photo analysis and retroactively binned photo analysis survey data were able to produce reliable regressions with $r^2$ values of 0.57 and 0.56, respectively. Overall, this study demonstrates that photo analysis surveys are the most reliable dataset to use when creating crop residue cover models, but we also acknowledge that these surveys are expensive to conduct and suggest some ways these surveys could be made more efficient in the future.

**Keywords:** crop residue cover; normalized difference tillage index (NDTI); tillage intensity; residue survey; crop residue indexes; Google Earth Engine

## 1. Introduction

Excessive soil erosion is a significant environmental issue that is particularly acute within agricultural landscapes. The topsoil losses within agricultural regions have been shown to be extensive, reducing soil productivity and resulting in severe economic impacts [1,2]. Sediment and nutrient runoff also have a negative impact on the downstream infrastructure [3] and ecology [4] that can be costly to remediate. Consequently, understanding the magnitude and patterns of soil erosion within agriculture landscapes is an important pursuit that will positively impact society and the environment.

Due to the complex interactions that govern sediment erosion and transport [5], monitoring erosion within large regions is challenging, and models that enable this require a variety of biophysical input parameters [6]. One model that has had success in this regard is the Daily Erosion Project (DEP), which generates daily estimates of precipitation-related soil loss for over 6000 small watersheds (USGS Hydrologic Unit Code 12) across the Midwestern United States [7]. The DEP accomplishes this via an implementation of

the Water and Erosion Prediction Project (WEPP) [8] model that utilizes radar-derived precipitation data and additional weather data, LiDAR-derived terrain data, field-level crop rotation data [9], and SSURGO soil property data [10] to ascertain most of the input parameters needed for the WEPP model. While the WEPP input parameters that are sourced from these data are relatively well constrained, information regarding the crop residue cover conditions of agricultural fields has proven to be difficult to obtain while maintaining accuracy, resolution, and coverage [11]. The DEP implementation of the WEPP infers one of six tillage regimes and estimates time-based changes in residue cover, soil bulk density, soil roughness, etc., based on the minimum observed residue at one or more time periods during the tillage calendar [8]. It is thus understood that the residue cover estimates within the DEP/WEPP may not match the observed residue cover by humans or sensors in some fields due to cloud coverage or model errors, but this was selected as the most practical method of estimating the tillage practices required by the WEPP and similar models like RUSLE2 and SWAT.

The tillage information needed to estimate crop residue cover and/or assign one of the six DEP tillage practices could potentially come from either tillage surveys or crop residue cover measurements. Tillage surveys, which are traditionally conducted by driving a transect across a geographic area such as a county and observing the tillage practices at a defined interval, were routinely conducted by the United States Natural Resource Conservation Service (NRCS) nationally from 1982 through 2004 as part of the Crop Residue Management by the Conservation Technology Information Center (CRM; CTIC) survey and are still conducted in some locations. The field-level data recorded while driving these transects are typically divided into residue-cover- and technique-based tillage classes as defined by the survey methodology (CTIC), such as conventional till (<15% at planting), reduced till (15–30%), and strip till, ridge till, and no till (all > 30%), potentially allowing one to estimate the mean residue cover observed for each tillage class. Alternatively, some surveys classify the field residue cover conditions into a 10% crop residue cover bin (e.g., 0–10%, 10–20% . . . 90–100%). Since this method does not accommodate the break at 15% that is found in the NRCS and CTIC conventional and reduced tillage system definitions, it is not utilized as often. Both methods are conducted via the "windshield", meaning the observer does not leave the vehicle, which enables quick collection of data, but these estimates often exhibit bias [12–14], which is likely due to the human difficulty in converting viewing angles from oblique to nadir [13], resulting in recommendations such as holding one or more observers consistent across multiple county surveys [15]. New techniques to standardize the windshield assessment process are being developed [16] but are not yet widely applied. Along with these "windshield" methods, in-field methods are also used to estimate crop residue cover. In-field crop residue cover measurements can be collected via a multitude of methods, including tape transects, spiked wheels, and point-count photo analysis surveys [17]. These methods all have inherent relative advantages and disadvantages but are generally more accurate than the "windshield" methods. However, the in-field techniques tend to achieve lower levels of adoption due to the larger amount of time and effort they take to complete.

The time constraints of observing the residue cover in millions of fields of 500,000+ km$^2$ necessitate the use remote sensing, which many recent studies have demonstrated the effectiveness of in terms of using single- or multi-band indexes within regression or machine learning models to predict crop residue cover [18–23].

To minimize the computational expenses and training data variability associated with increasing geographic extent, a problem not unique to residue cover studies [24], an important recent development is the creation of cloud-based geospatial analysis platforms such as the Google Earth Engine [25]. The Google Earth Engine hosts all the major publicly available Earth observation satellite datasets as well as some aerial datasets (e.g., US National Agricultural Imagery Project; NAIP) and provides an optimized computational environment that greatly reduces computation times when compared to running similar analyses on standard desktop machines. The Google Earth Engine introduces some level of

image abstraction, which can result in the loss of insight into the processing steps, such as easily identifying which image date within a time range contains specific pixel values. The Google Earth Engine can also help to overcome the difficulties associated with the temporal and spatial variations in soils and crop residue cover over large regions by incorporating remote sensing derivatives that are sensitive to changes in the photosynthetic vegetation and soil moisture content within the analysis or removing problematic data, as was shown before [21,26], thus expanding the training data coverage to multiple non-contiguous regions over multiple years.

In this study, our primary objective was to test multiple sources of training data and determine the minimum quality of crop residue fraction that is needed to produce reliable crop residue cover models via the creation and comparison of models trained using residue cover measurements collected using a variety of methods. Our secondary objective was to test the ability of different crop residue cover data sources to identify the endpoints of the residue cover spectrum. Between 2010 and 2017, the authors and cooperators collected crop residue survey data using three different survey techniques (windshield tillage surveys, windshield binned residue surveys, and photo analysis surveys) across three Midwestern US states where corn and soybeans are commonly grown. These survey data were used in conjunction with a Google Earth Engine spatial analysis and a local linear regression analysis to assess whether each survey type is capable of producing reliable regression models. Additionally, we tested the utility of photo analysis survey data that were retroactively binned into 10% residue cover bins. We hope that this study will help future researchers to select the crop residue cover data collection strategy that is most appropriate for their study design when creating remote-sensing-based residue cover models.

## 2. Materials and Methods

### 2.1. Data Collection

To investigate the suitability of windshield tillage surveys, windshield binned residue surveys, and photographic analysis surveys as sources of training data for a crop residue cover model, we and our colleagues have collected and aggregated residue cover estimates via various survey methods throughout the Midwestern United States. All survey data were collected during the springtime between 2010 and 2017 within the states of Nebraska, Kansas, and Iowa, USA. Table 1 provides a summary of collection years and states for each survey dataset and Figure 1. provides a map of the survey locations. Iowa data were collected using two different survey techniques depending on the year and location surveyed, while Nebraska and Kansas survey data were collected using consistent regional surveys.

**Table 1.** Collection method, state, years, and number of fields for all survey data within our study.

| Survey Type | State | Years | Number of Fields |
|---|---|---|---|
| Windshield Tillage | Nebraska | 2016 | 5561 |
| Windshield Tillage | Kansas | 2010, 2013, 2014, 2016 | 4667 |
| Windshield Binned Residue | Iowa | 2013, 2017 | 1615 |
| Photographic Analysis | Iowa | 2015, 2016, 2017 | 296 |

The Iowa residue cover surveys were conducted using both windshield binned surveys and photo analysis surveys between 2013 and 2017 within four different survey campaigns. The 2013 windshield binned residue cover survey was conducted by Agren, Inc. (Carroll, IA, USA), an agricultural conservation planning business, during the month of May via observations of residue cover in all fields within the South Fork of the Iowa River, an HUC8 watershed in north central Iowa. Observations were collected by trained personnel and classified into 10% residue category bins (e.g., 0–10%, 10–20%, . . ., 90–100%). The 2016 binned residue cover survey was conducted by watershed coordinators as part of the Iowa Watershed Approach flood management project. This data collection was coordinated by the Iowa Department of Natural Resources using personnel with similar backgrounds

and similar procedures regarding the South Fork data from 2013. Photo analysis surveys were conducted in 2015, 2016, and 2017 after planting (late April through early June) using the reference methods described by Gelder et al. [19]. All field photos were taken from approximately 1 m directly above the field (Figure 2), and four to five images were collected for each field at random locations at least 30 m apart throughout the field. Trained observers placed grids over the images on a computer screen and the presence or absence of crop residue was noted for 100 evenly spaced grid intersections. Crop residue cover for each image was defined as the percent of intersections that corresponded to the presence of crop residue, and field crop residue cover value was defined as the mean crop residue cover of all the images from that field.



**Figure 1.** Summary map of survey data collection. Windshield tillage surveys were conducted in Nebraska and Kansas, while windshield binned residue and photo analysis surveys were conducted in Iowa. State names are located in the bottom left-hand corner of the states, and the inset map shows the study states within the contiguous United States.

Nebraska surveys were conducted using the CRM windshield tillage technique and were conducted in late April through early June 2016 using the methods developed by the Nebraska state office of the NRCS. These surveys were conducted via observing the type of tillage found within fields at one-mile intervals along a predefined route. For both the Nebraska and Kansas windshield tillage surveys, the tillage practice present in each field was recorded as conventional tillage (<15% residue coverage), reduced tillage (15–30% residue coverage), conservation tillage (30–50% residue coverage), or no till (>50% residue coverage). This information was collected using a GPS-enabled tablet to assist in accurate geolocation during data entry.

**Figure 2.** Two examples of crop residue images used in our study. The top image shows a field with 11% soybean residue cover with some emergent corn, and the bottom image shows a field with 79% corn residue cover. All photos were taken from one meter above the field, and overlaid grids, such as the ones shown in the right half of the bottom image, were used to estimate the spatial coverage of crop residue within the images. Horizontally, gridlines are 7.62 cm (3 inches) apart.

The Kansas surveys were also windshield tillage surveys and were conducted by a team of 3–4 trained personnel who drove a pre-determined route through each county, stopping at intersections and determining the tillage practice present in each field. Approximately 450 fields per county per year were examined, both in fall and again in spring, but only the spring data were used in our analysis.

Before the Kansas and Nebraska survey data could be used to form regression relationships, the tillage categories needed to be converted to numeric crop residue cover

values. To facilitate this, the observed tillage type and the Cropland Data Layer (CDL) crop type from the previous year were used to convert tillage practice observations to estimates of spring crop residue cover using the conversions of Bull [27] (Table 2). The windshield tillage surveys were performed using these converted crop residue cover values.

**Table 2.** Values were used to convert tillage type observations and CDL-based crop type combinations to spring crop residue cover for the windshield tillage survey data. This conversion was performed so that tillage survey data could be used to create regression relationships. The conversions used in this study are taken from Bull (1993).

| Tillage Type | Crop Type | Converted Residue Cover Value |
|---|---|---|
| Conventional Till | Soybeans | 3% |
| Conventional Till | Corn | 4% |
| Reduced Till | Soybeans | 26% |
| Reduced Till | Corn | 34% |
| Ridge Till | Soybeans | 42% |
| Ridge Till | Corn | 55% |
| Strip Till | Soybeans | 42% |
| Strip Till | Corn | 55% |
| Mulch Till | Soybeans | 55% |
| Mulch Till | Corn | 72% |
| No Till | Soybeans | 55% |
| No Till | Corn | 72% |

In addition to these surveys, the photo analysis survey data were artificially binned into 10% bins (e.g., 0–10%, 10–20%, ..., 90–100%) to create a binned photo analysis dataset. This dataset was used to determine if photo analysis residue cover data classified to a coarser resolution could be usable as training data.

### 2.2. Geospatial Analysis

For the geospatial analysis, residue cover survey data were associated with field polygons on a local computer before they were uploaded to Google Earth Engine, where the satellite imagery analysis was conducted. A tabular version of the results from this analysis was downloaded to a local computer, where the final regression analysis was conducted within a Python 3.4 computational environment. A graphical summary of our analysis workflow can be found in Figure 3, and all Google Earth Engine/Python code used in this study can be found at the GitHub repository: https://github.com/dailyerosion/Williams_et_al_2024_residue_cover (accessed 19 August 2024). An in-depth description of our workflow follows.

**Table 3.** Satellite and index information for the indexes used within our study. For each Landsat satellite platform, we used NDVI and NDMI to remove image data associated with excess emergent vegetation and excess soil moisture, respectively. Then, we used NDTI to estimate crop residue cover on quality raster cells.

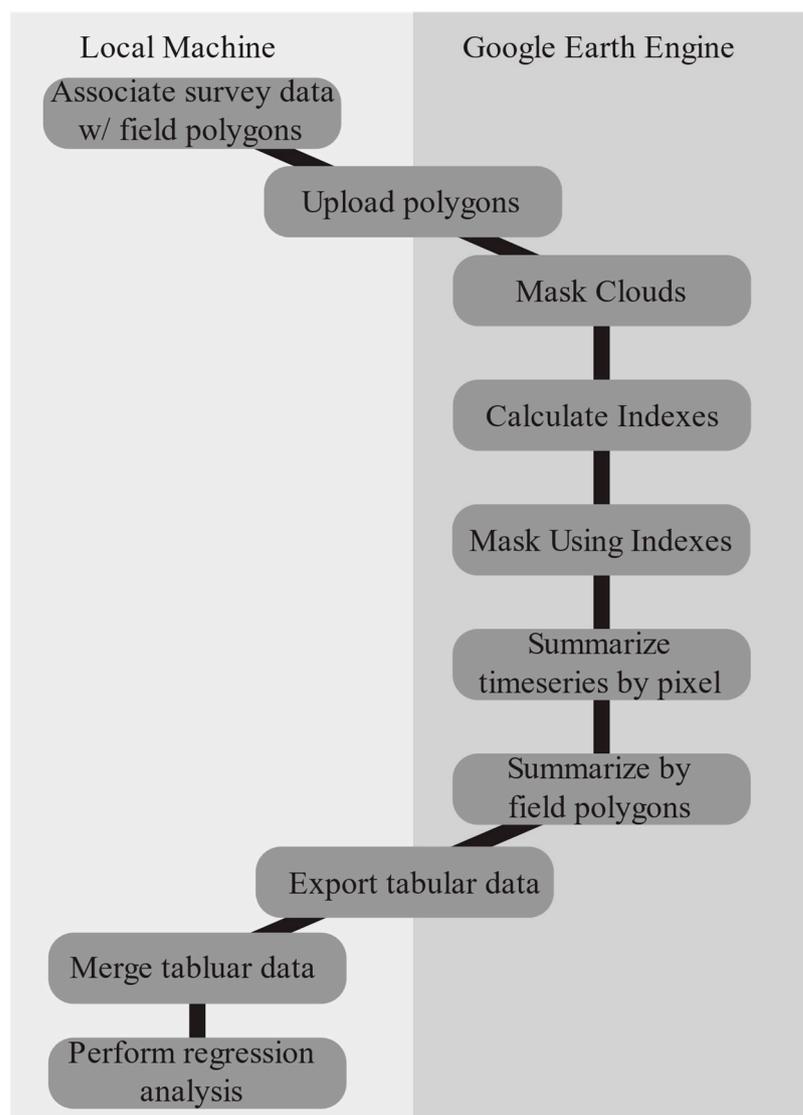| Platform | Ground Resolution | Availability | Processing Level | Name | Equation |
|---|---|---|---|---|---|
| Landsat 7 | $30 \times 30$ m | 1999–2022 | SR | NDTI | $\frac{B5-B7}{B5+B7}$ |
| Landsat 7 | $30 \times 30$ m | 1999–2022 | SR | NDVI | $\frac{B4-B3}{B4+B3}$ |
| Landsat7 | $30 \times 30$ m | 1999–2022 | SR | NDMI | $\frac{B4-B5}{B4+B5}$ |
| Landsat 8 | $30 \times 30$ m | 2013–Present | SR | NDTI | $\frac{B6-B7}{B6+B7}$ |
| Landsat 8 | $30 \times 30$ m | 2013–Present | SR | NDVI | $\frac{B5-B4}{B5+B4}$ |
| Landsat 8 | $30 \times 30$ m | 2013–Present | SR | NDMI | $\frac{B5-B6}{B5+B6}$ |

**Figure 3.** Workflow for our analysis. Field data were associated with field polygons on a local machine, then uploaded to Google Earth Engine where we associated each field polygon with the indexes listed in Table 3. A tabular version of these data was then downloaded to a local computer where we performed our regression analysis.

Residue cover photographic survey data were first recorded as pointwise observations. However, since crop residue cover is likely to vary most significantly at the agricultural field scale, each point was associated with the appropriate agricultural field polygon sourced from the Agricultural Conservation Planning Framework database [9]. This was conducted so that the analysis could be conducted on the field scale for all analyses since all windshield data were only collected as a field average. In the case of the photo analysis surveys, the mean value of all points within a field was used as the field's residue cover value. All field polygons were buffered inward by 30 meters to limit the influence of edge effects, then uploaded to Google Earth Engine where the satellite analysis was conducted.

Within Google Earth Engine, time series data from the two most relevant Landsat Earth observation missions (Landsat 7 and 8) were used to calculate spectral indexes that relate to the three key components of residue cover observations: the amount of residue cover, emergent photosynthetic vegetation, and soil moisture. To capture these properties, we utilized the Normalized Difference Tillage Index (NDTI), Normalized Difference Vegetation Index (NDVI), and the Normalized Difference Moisture Index (NDMI). The NDTI focuses primarily on soil surface differences. This is essential for calculating the residue cover

percentages. The NDVI and NDMI are primarily included and used for filtering out different pixels for being too green or too wet. Those two ensure that the final pixels being reviewed and analyzed are the most likely to have the residue that is being searched for. The equations used to calculate these indexes can be found in Table 3. Using the pixel quality bands that are a component of the Landsat 7 and Landsat 8 data, bitwise masks were used to remove pixels containing clouds or snow prior to analysis. For Landsat 7, bits 3, 5, and 7 were used, and for Landsat 8 bits 3 and 5 were used. Once clouds and snow were removed, the Landsat 7 and 8 surface reflectance datasets were harmonized using the methodology of Roy et al. [28], and the resulting dataset was used to calculate our spectral indices. Both Landsat 7 and 8 mission data have ground sampling distances (GSDs) of $30 \times 30$ m.

To capture all possible images that could correspond to the time of spring planting, when the DEP model expects minimum residue cover to be measured [7], all images that were collected between 1 March and 15 June each year were included. On average, this results in 6 cloud-free images for Landsat 7 and 8 combined, but this varied based on year and location (Table 4). It was not possible for us to determine the time of planting at each pixel, so instead we used the minimum NDTI value of each year's image stack to produce the NDTI value used in our regression analysis. The minimum NDTI was selected based on previous work by Gelder et al. [19] and Zheng et al. [14], which both showed that using the minimum value of the image stack accurately corresponds to the residue cover at the time of planting.

**Table 4.** Average number of valid images per field within the analysis dates (1 March and 15 June). Valid images are defined as images that pass the NDVI/NDMI thresholds described above and are free from cloud cover. This combines imagery from both Landsat 7 and 8, resulting in higher numbers of valid images when both satellites were operational (2013+).

| Survey Type | Year | Avg. Valid Images |
|---|---|---|
| Windshield Tillage | 2010 | 2.4 |
| | 2011 | 1.6 |
| | 2014 | 9.1 |
| | 2015 | 4.1 |
| | 2016 | 5.9 |
| | 2017 | 5.9 |
| Windshield Binned | 2013 | 5.9 |
| | 2017 | 6.7 |
| Photo Analysis | 2015 | 4.9 |
| | 2016 | 7.1 |
| | 2017 | 3.6 |

Similarly to Beeson et al. [29], we controlled for extreme NDTI values by removing NDTI values associated with high levels of emergent vegetation and/or excess soil moisture. In keeping with Beeson et al., we used an NDVI threshold to remove pixels overly influenced by emergent vegetation, but unlike their study we used an NDMI threshold instead of NEXRAD rainfall data to remove pixels with high soil moisture. While this is a slightly different metric, we expect that our NDMI threshold approach will behave similarly to their NEXRAD approach and has the advantage of exclusively relying on Landsat data.

To determine the most appropriate threshold values, we tested different NDVI/NDMI thresholds and determined the optimal threshold values based on model performance and data coverage. We calculated the average NDVI and NDMI index values within each field for each Landsat image available within our sensing period. We then used these values to create distributions for each index (Figures 4 and 5). Using these distributions, we assessed the utility of using the 50%, 60%, 70%, 80%, and 90% quantile values for the NDVI/NDMI index threshold values. We decided that using an 80% quantile threshold provided the best balance between coverage and data quality, and we used this threshold value for both NDVI and NDMI. These 80% threshold values were 0.129 and −0.081 for NDMI and NDVI,

respectively. In addition to using these thresholds to discard low-quality data, any raster cells that did not have data available from three valid Landsat images after thresholding were also discarded, ensuring that the minimum NDTI calculation is based on at least three image captures.
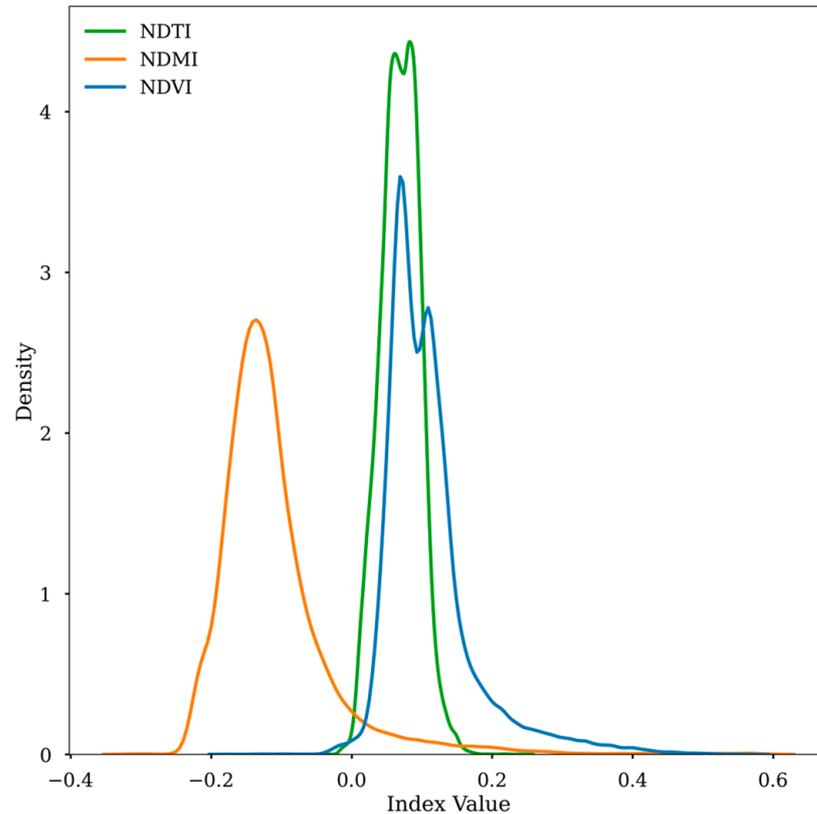


**Figure 4.** Normalized kernel density estimated distributions for each of the three indexes used within the study period. NDTI is the most normal, and the NDMI/NDVI distributions are slightly right-skewed.

The use of SSURGO soil survey data [30] as a regression factor was considered but not implemented due to multiple reasons. Soil color has been shown to have an impact on residue cover determination [20,31], but SSURGO data are not currently available as a Google Earth Engine dataset, and others have shown that SSURGO texture data may not improve model performance in a remote sensing context [29]. Additionally, SSURGO mapping units were designed with minimum areas of 1–5 ha, and complexes (mixed soil series) are mapped in areas where soil series are not readily separated [32]; thus, they were not designed to be accurate at the GSD of current Landsat images.

Once time series values were calculated at the pixel level, they were then aggregated to the field level by taking the median summarized index value within the inward-buffered field polygon. Aggregating to the field level by taking the minimum field index value was not considered because this would likely include data from in-field depressions whose high soil water content biases NDTI and NDVI values [19,20] and mean values would likely include high-residue-cover waterways or other conservation practices. In addition to the indexes described above, the modal Cropland Data Layer (CDL) value for each field during the previous year was used to assign a crop type to each field.
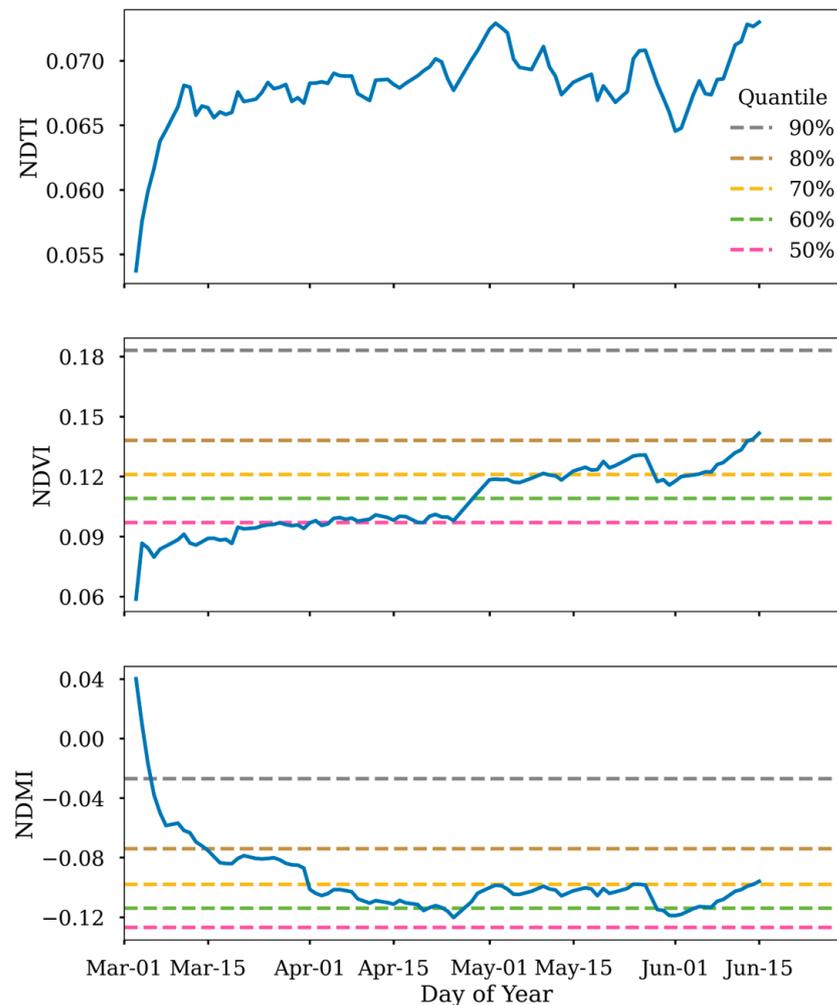
**Figure 5.** Thirty-day rolling average for each index and day of year and quantiles of NDVI and NDMI for thresholding. NDTI and NDVI display similar trends of increasing as the year progresses. In contrast, NDMI appears to reach a stable minimum in April.

### 2.3. Regression Analysis

While other crop types were present in small numbers within our training data, we chose to focus our analysis on corn (*Zea mays*) and soybean (*Glycine max*) fields, which composed 52% and 42% of our survey data, respectively. The remaining 6% of our data, which was composed of 17 different crop types, was discarded. Even though corn and soybean crop residues have different structures (Figure 2), we found little difference in their NDTI response for a given crop residue cover percentage in initial trials, and thus performed the regression for both crops at once. This may be due to the fact that long-term low-tillage-intensity fields in corn–soybean rotations have significant corn residue even in the year directly after soybeans.

To assess the quality of the data produced by the various survey techniques, we conducted four separate linear regression analyses between each of the four survey types used and the satellite-data-derived NDTI values. We then used the quality of these regressions to determine the utility of using each survey type as training data for a remote-sensing-based residue cover model. We assessed model fit for all regressions via the coefficient of determination ($r^2$).

We are also interested in the feasibility of using each type of training data for discriminating between the end-member tillage types (e.g., differentiating between no till and conventional till, corresponding to maximal or minimal residue cover) and thus performed a secondary analysis to assess each survey type's utility in this context. For this analysis, we

removed data not associated with either no till or conventional till fields from each dataset, then compared the distributions of each dataset's NDTI values between the two remaining tillage types. For the windshield tillage dataset, the original tillage type classifications were used to distinguish between no till and conventional till fields. For the remaining datasets, any field with less than 10% observed residue cover was classified as a conventional till field and any field with greater than 55% observed residue cover was classified as a no till field. The resulting distributions were assessed based on their degree of separation using the Kolmogorov–Smirnov test, a non-parametric test of the dissimilarity between two distributions.

## 3. Results

### 3.1. Temporal Index Trends

Figure 4 displays the distributions of the NDTI, NDVI, and NDMI indexes throughout the study period. The NDTI and NDVI both had mean values close to zero (0.05 and 0.10, respectively, while the NDMI had a lower average value of $-0.12$. The NDTI distribution was roughly normal, but the NDVI and NDMI distributions were slightly right-skewed, indicating that some areas in the study area have consistently high soil moisture or green vegetation. A 30-day rolling average of the mean value for each index can be found in Figure 5. In general, the NDMI values are highest in the early spring when snow and snowmelt are present. Conversely, the NDVI steadily increases throughout the sensing period and likely continues to increase as the growing season progresses. Notably, the NDTI values show more variability than either the NDVI or NDMI values and also seem to follow a similar trend in late spring to the NDVI values. In particular, the sharp increase in the NDTI values after June 1st mirrors a similar increase in the NDVI values around this time. This indicates that the increases in the NDTI values around this time are likely related to new green crop growth, agreeing with Gelder et al. [19].

### 3.2. Survey Data Model Performance

The regression performance between the windshield-based and field-based surveys exhibited stark differences. In the windshield tillage survey, no clear visual relationship between tillage type and NDTI values was observed (Figure 6), resulting in a coefficient of determination ($r^2$) of 0.21. This indicates a weak correlation between the two variables. Similarly, the windshield binned survey exhibited a limited relationship between the observed residue cover and NDTI values, yielding a slightly improved $r^2$ of 0.24. However, it was noted that lower NDTI values were more frequently associated with low observed residue cover.

In contrast, the photo analysis survey demonstrated a significantly clearer relationship between the observed residue cover and NDTI values. Notably, a distinct linear trend was observed when the residue cover exceeded 30%. However, below this threshold (Figure 6), a higher incidence of outliers was observed, and no clear trend emerged. The overall relationship exhibited an $r^2$ of 0.56 when considering all the data points. When restricting the analysis to data points with residue cover values greater than 30%, the $r^2$ improved to 0.64, indicating a moderate-to-strong correlation.

The retroactively binned photo analysis survey data exhibited a performance similar to the non-binned version (Figure 6). The regression analysis resulted in an $r^2$ of 0.56, indicating a moderate correlation between the observed residue cover and NDTI values. Similar to the non-binned model, the model performed less effectively in predicting residue covers below 30%. Notably, the regression models for both versions of the data were nearly identical. For example, for an NDTI value of 0.05, the non-binned model estimated a residue cover of 54.3%, while the binned model estimated 54.0%, a marginal difference of 0.3%. These findings emphasize the superiority of photo analysis surveys over windshield-based surveys in capturing the relationship between residue cover and NDTI values. The photo analysis approach exhibited a clearer and more consistent trend, particularly when the residue cover exceeded 30%.
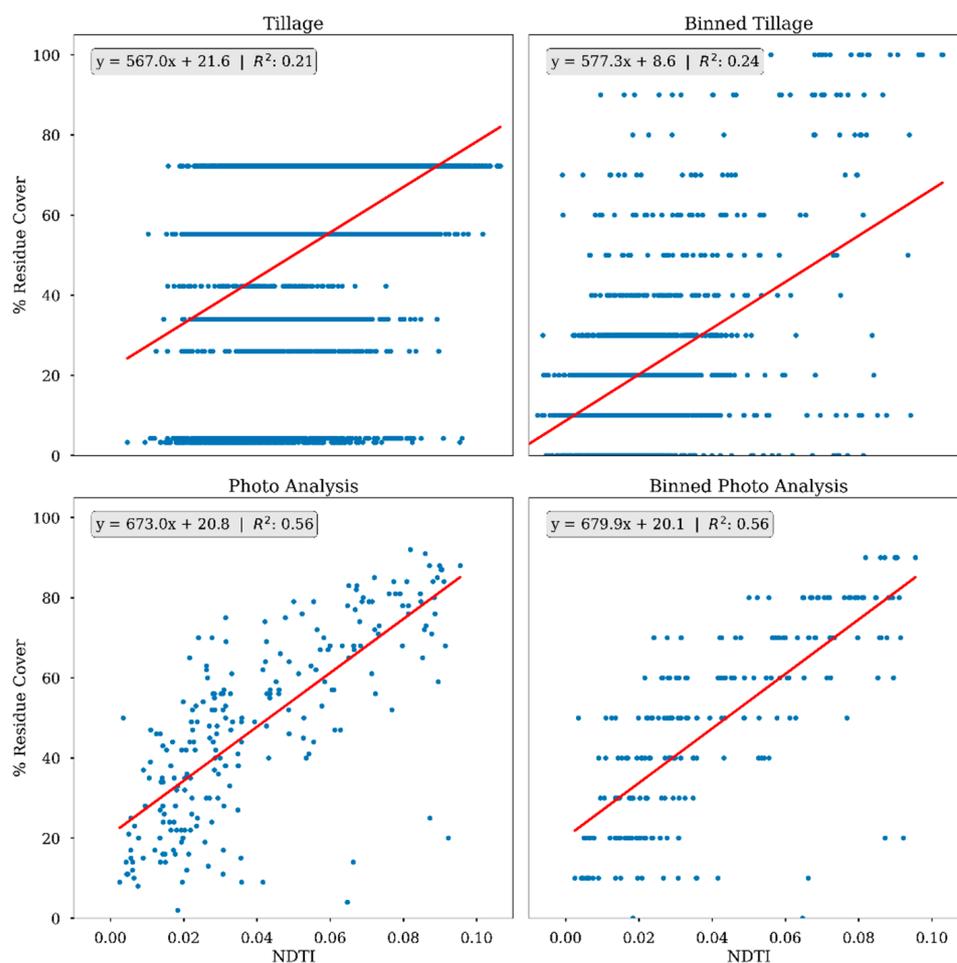
**Figure 6.** Linear regression models for survey datasets. Both the windshield tillage and windshield binned surveys performed poorly, but the photo analysis and the retroactively binned photo analysis produced much better results.

### 3.3. End-Member Analysis

While our windshield tillage survey data produced poor regression results, when looking at the end-member distribution of the minimum NDTI values for the "no till" and "conventional till" tillage groups, we do see a clearer separation in the distributions (Figure 7), with a K–S value of 0.50 and a *p* value of $2.2 \times 10^{-129}$. This indicates significant differences in the minimum NDTI distributions for the observed "no till" and "conventional till" fields. All the other surveys also showed significant differences between the minimum NDTI distributions for the two end-member regimes, with a concomitant increase in the K–S statistic that correlated to increases in $R^2$ within the linear regression analysis. This indicates that all the survey types are generally able to distinguish between "no till" and "conventional till", although the photo surveys distinguish the end members better. This consistent trend of NDTI increase across all the surveys as residue cover increases within similar ranges indicates that these models are generally suitable for application across large areas with accurate ground truth data. The photo analysis survey data had the highest K–S statistic value (0.81), followed closely by the binned photo analysis survey data. Notably, the *p* values are smaller for the windshield tillage survey types due to the larger number of samples available for these surveys.
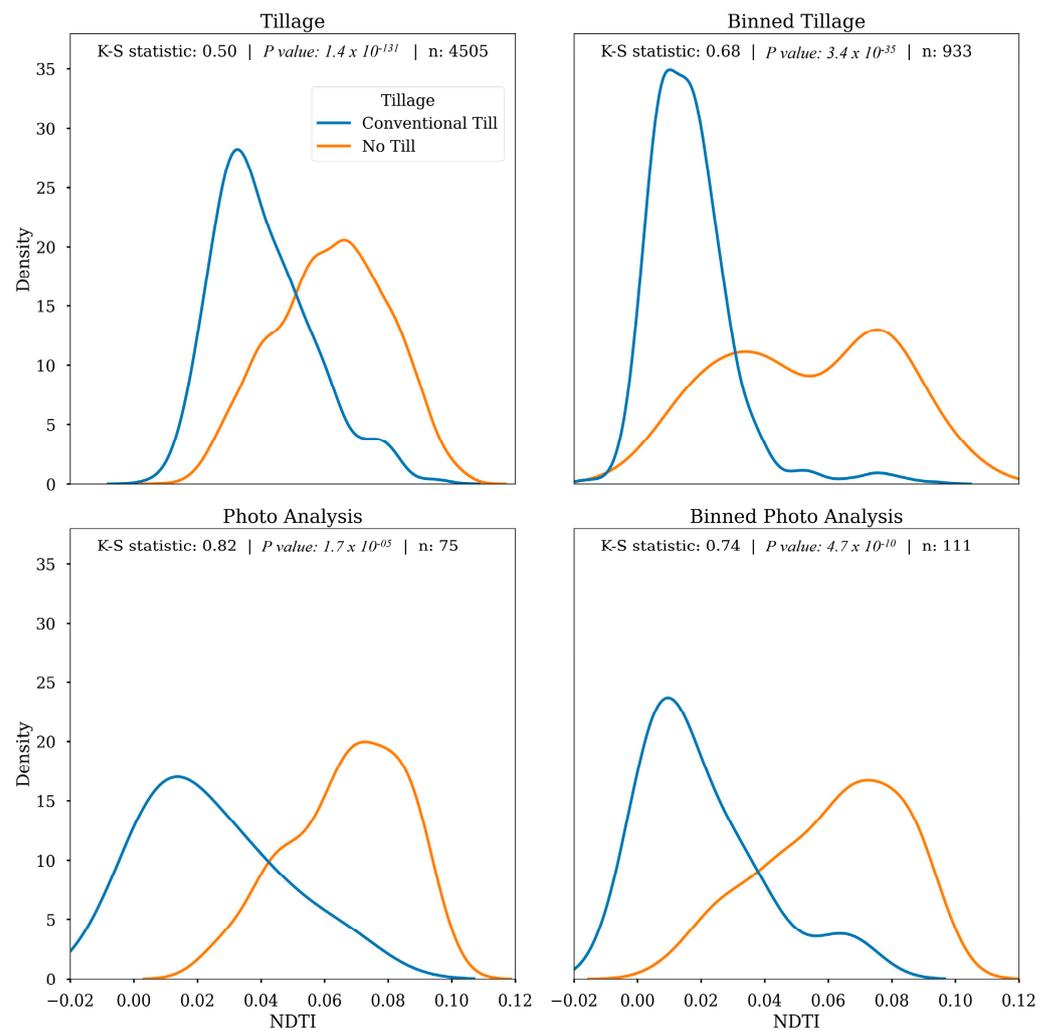
**Figure 7.** Distribution of minimum NDTI values within the two main tillage regimes for all survey types, and for windshield binned residue data. There is a large amount of overlap between the top two distributions for windshield tillage survey data, whereas the separation is more apparent for photo analysis survey data, confirmed by higher K–S statistics.

*3.4. Application to Regional Crop Residue Cover Analyses*

Based on our regression analysis, photo analysis survey data created the best regression when paired with Landsat 7/8 minimum NDTI values and NDVI and NDMI thresholds. The resulting linear model used to predict crop residue cover is as follows:

$$Crop\ Residue\ Cover\ (\%) = 676 \times Landsat\ 7\&8\ NDTI + 20.5$$

The crop residue cover map produced using this equation for 2021 can be found in Figure 8. Overall, our predictions of crop residue cover align well with the USDA's Major Land Resource Areas (MLRAs). MLRAs are geographically associated land resource units that contain similar types of soils, climate, and land use and thus have generally similar tillage practices, suggesting that our model is capturing the broad variations in tillage management that occur throughout the state.
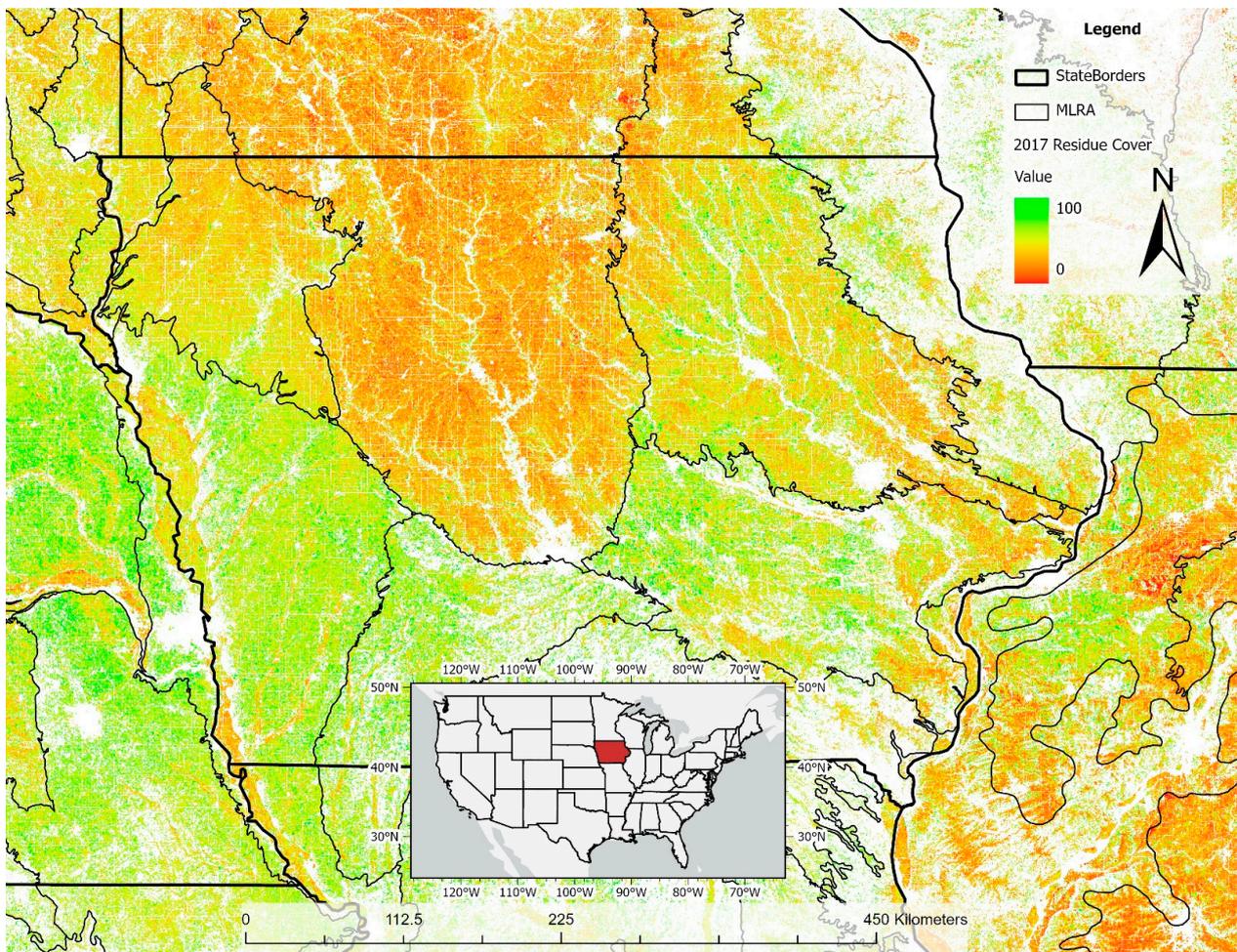
**Figure 8.** Percent crop residue cover predictions produced for 2017 using the optimal regressions from our analysis. Variations in predicted crop residue cover align with Major Land Resource Areas (MLRAs), suggesting that our model is capturing the broad variations in tillage management that occur due to soils and crops. Inset map shows the location of the AOI within the coterminous United States.

## 4. Discussion

### 4.1. Recommended Use Cases of "Windshield" Survey Datasets

While windshield tillage and windshield binned residue surveys are widely conducted, as evidenced by the large number of observations we were able to assemble, our analysis shows that they perform poorly when used to predict crop residue cover as a continuous variable. This is disappointing because datasets such as these are widely available throughout the United States, have existed for many years, and thus could be a valuable training dataset. Both windshield survey datasets had high levels of variability, which led to training data that were distributed over the entire range of predicted crop residue cover values for a single tillage class or residue cover bin, making it difficult for the models to identify the relationship observed in the photo analysis dataset.

Although neither windshield dataset produced reliable regressions, the distributions in Figure 7 show that binned residue data can generally be grouped into "low intensity" and "high intensity" tillage categories that have distinct minimum NDTI value distributions, although the photo analysis data have less overlap, which we hypothesize is due to the improved observation angle [13]. The regression results were not unexpected because they are likely due to the inherently inaccurate data conversion process ([27]; Bull et al.) between categorical information and quantitative data. This separability of low/high-intensity tillage suggests that windshield data may likely be useful in an end-member

tillage classification approach if the quality can be maintained. We do not believe that this contradicts the findings by others [13,22] who found that some high-quality tillage class training datasets could be used to develop models that accurately estimate the tillage classes on validation fields. This finding is particularly relevant to crop residue cover approaches that employ spectral unmixing or otherwise require information on the end-member characteristics of tilled fields [29,33]. In this context, windshield binned residue surveys should be able to reliably represent the end-member characteristics of high-tillage and low-tillage fields without the need to employ more rigorous survey techniques. However, actual residue cover values, and not the standard three or four tillage classes, are preferred by the DEP due to the improved ability to discriminate between the six DEP tillage classes that may only differ by a few percent in terms of the residue cover at planting in the more intense tillage systems.

*4.2. Recommended Use Cases of Photo Analysis Survey Datasets*

As expected, the photo analysis survey datasets produced the most reliable models of crop residue cover. Unlike the windshield methodologies, the photo analysis survey observes the field from above, samples multiple portions of a given field, and quantitatively assesses the amount of crop residue cover in each location. All these factors combine to create a more accurate representation of in-field crop residue cover, and, consequently, this technique returned the best correlations with the remote sensing estimates. As noted in the Introduction, however, this is also the most labor-intensive technique, so the number of data points collected to establish the crop residue models will also need to be balanced against the time and monetary expenses of this approach. Alternative photo collection and photo analysis techniques that maintain the same level of data quality but decrease the cost of collection will likely need to be developed before we observe the widespread adoption of this survey technique.

In contrast to the windshield data, the retroactively binned photo analysis survey produced a regression that performed almost equally well compared to the photo analysis survey data. This finding suggests two things. First, the issues found within the windshield binned residue data do not inherently result from binning data, and efforts to improve the accuracy of windshield binned residue surveys [16] may make them suitable for use within a regression framework. Second, those photo analysis techniques that increase the percent crop residue cover speed for a given image that also sacrifice a small amount of accuracy may be worth pursuing. For example, in our analysis of photos, we recorded the presence or absence of crop residue in 100 points per image. Based on these results, it is likely that it would be possible to reduce the number of points observed per image significantly without reducing the efficacy of the resulting crop residue cover model, which would in turn enable us to collect more training data in the same amount of time. Additionally, machine vision approaches [34] that can reliably classify residue cover images into 10% bins could offer a way to quickly collect large quantities of reliable training data.

*4.3. Model Limitations*

The best-performing model was able to produce linear regression with an $r^2$ of 0.56, which resulted in residue cover maps that correlate with the known variances in crop residue cover percentages across Iowa. While this article covers the effects of the input survey data on the resulting model, the quality and quantity of the available remote sensing data are also important factors that influence these models. For example, it is well documented that local variations in confounding factors such as soil moisture, emergent vegetation, and soil type/color can have a significant impact on the accuracy of the resulting remote sensing models [35–37], which could explain the reduced correlation between the NDTI and residue cover below 30% in Figure 6; thus, accuracy should be assessed when estimating residue cover on soil orders not sampled in this study. The NDTI values in Figure 7 across all the survey types indicate similar minimums and maximums across the Iowa, Kansas, and Nebraska soils, indicating that the regressions across the entire

region are generally valid but may have decreased accuracy when compared to locally derived regressions. Generally, we succeeded in reducing the impact of the two most important confounding factors (soil moisture and emergent vegetation) for estimating residue cover; however, better characterization of these factors would likely result in model improvements.

In addition, since we are using minimum spring NDTI values to predict crop residue cover and requiring a minimum of three valid images, we are limited by the number of remote sensing observations we have in a given location. As shown in Table 4, we have on average five valid image dates between 1 March and 15 June; thus, some locations will not have enough observations in a given year to generate an estimate. To overcome this, several possibilities exist. First, we could incorporate data from a wider variety of satellite platforms. In particular, the Harmonized Landsat Sentinel-2 dataset is promising for our work, and we look forward to the full development of the Sentinel 2 component within the Google Earth Engine. Second, we could utilize estimates from prior or subsequent years and assume similar practices existed in those years. Third, the use of alternative satellite technologies such as Synthetic Aperture Radar (SAR) [23] singly, or in conjunction with optical remote sensing, could expand the image opportunities during periods of cloud cover. In particular, the upcoming launch of the NISAR SAR satellite may provide interesting possibilities since it will be collecting fully polarimetric SAR data over our study region.

Although generally accurate in terms of determining the mean field residue cover, we observe in Figure 9, a detail of Figure 8 in central Hamilton County, Iowa, that the current model overestimates the range of field-level residue variation. The field to the south of the farmstead has significant soil organic and soil moisture content variation, and the model overestimates the variation in residue cover. The model-estimated field ranges (maximum–minimum) were roughly 30–40% Y, and the observed field range was 20%. We believe this is due to the linear regression model minimizing the errors for the mean field residue cover versus mean NDTI. We theorize that future research could improve the regressions by considering the soil property variation within each field and how that soil variation modifies the NDTI response to residue cover. This is likely to improve the low-residue-cover performance of the model as the soil background dominates the reflected spectra instead of the more spatially uniform residue signature.

### 4.4. Potential Cost-Effective Survey Designs

The major limitation of photo analysis surveys is the cost of transporting the survey staff to a diverse set of geographic locations to perform surveys and the cost of performing the point count photo analyses for the collected survey images. However, there may be ways to reduce both costs to the point that it is efficient to conduct photo analysis surveys at broad scales. First, the prevalence of GPS-enabled smartphones with high-quality camera systems means there are essentially no additional equipment costs associated with performing these surveys. With minimal training that could consist of a single instructional video, conservation groups or citizen scientists could be trained and asked to perform these surveys within their local area. They could then send these survey photos back to a centralized organization such as a university or conservation organization where the point-count surveys could be centrally performed. This approach could dramatically reduce the cost of obtaining raw survey photos and result in the collection of much more data.

In addition, our retroactively binned photo analysis results suggest that it may be reasonable to employ photo analysis techniques that speed up the process of measuring the crop residue cover percentages in survey photos, even if they sacrifice some accuracy. In the past few years, the ability of machine learning models to perform these types of analyses has improved rapidly [34], and using acceptable models would dramatically increase the speed of the photo analysis process. Thus, by combining distributed partner-led survey campaigns and the utilization of machine learning models for survey photo analysis,

the efficiency of these photo surveys could approach the efficiency of the conventional windshield surveys while providing much better data.
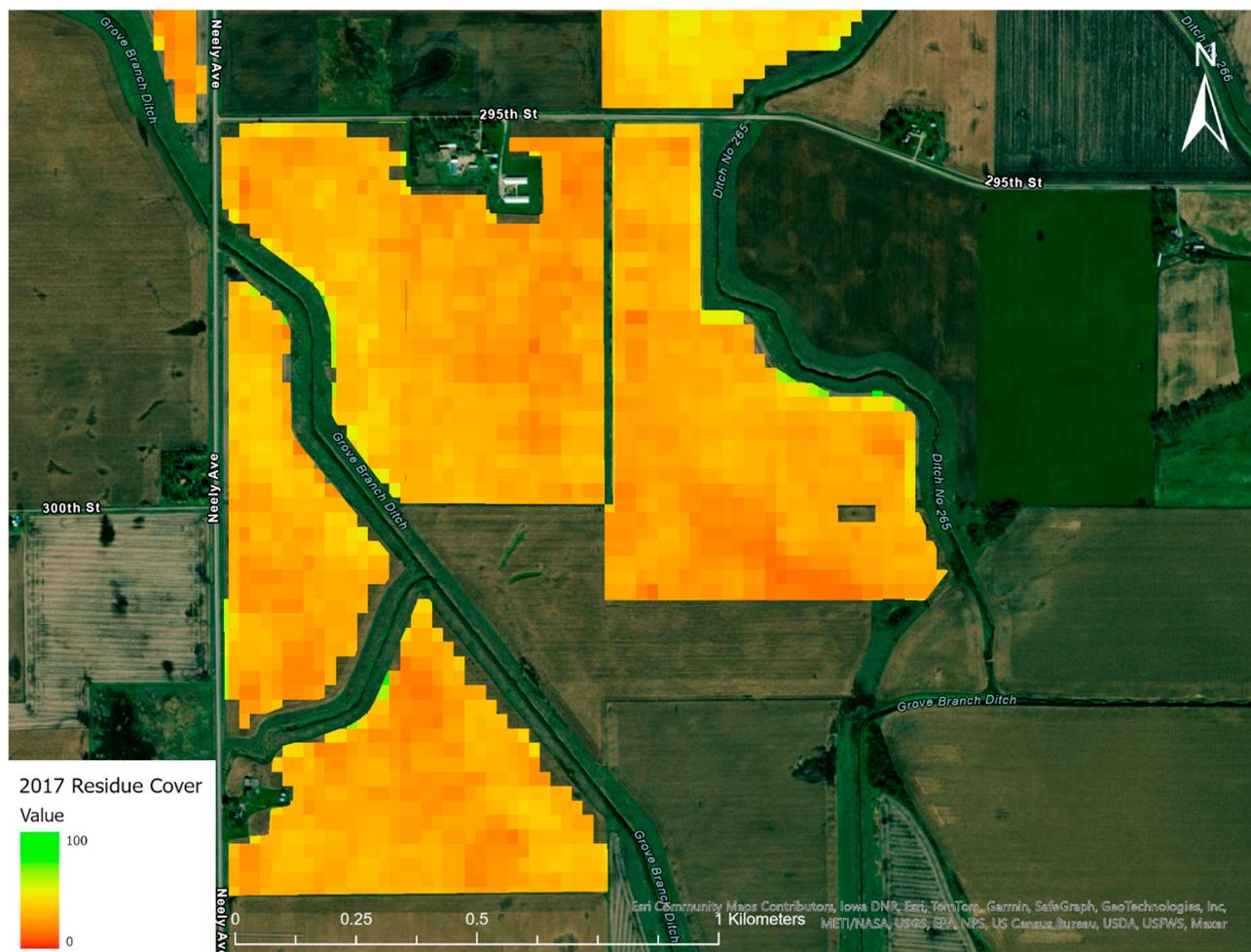


**Figure 9.** An example of specific in-field residue cover estimation using several fields located within Hamilton County, Iowa. Showcases overestimation of the field level for the range of residue being estimated.

## 5. Conclusions

In this study, we assessed the utility of three crop residue survey types (windshield tillage surveys, windshield binned residue surveys, and photo analysis surveys) and one synthetic survey (retroactively binned photo analysis data) as sources of training data for crop residue cover regressions. Using our highest-quality training data (photo analysis surveys), we were able to produce reliable estimates of crop residue for corn and soybean fields, but neither windshield-based survey dataset was able to produce reliable regressions. However, both windshield methodologies did produce a reliable distinction between low-residue and high-residue fields in aggregate, which may be useful in cases where the end-member characteristics of tilled fields need to be characterized. Finally, we found that retroactively binned photo analysis survey data performed equally as well as the original dataset, which suggests that more labor-efficient methods of retrieving crop residue cover from images should also produce reliable training data. The large-area implementation of this model in Figure 8 reflects the expected trends in residue cover. Further validation across time could verify accurate residue cover estimation across the entire Landsat 4 to Landsat 9 record, enabling the estimation of long-term tillage management information. However, this approach still exhibits inaccuracies within the field, over- and under-estimating residue cover on bright and dark soils. Overall, this study demonstrates that photo analysis surveys

are the most reliable dataset to use when creating crop residue cover models, but we also acknowledge that these surveys are expensive to conduct and believe these surveys could be made more efficient in the future.

## References

1. Pimentel, D.; Harvey, C.; Resosudarmo, P.; Sinclair, K.; Kurz, D.; McNair, M.; Crist, S.; Shpritz, L.; Fitton, L.; Saffouri, R.; et al. Environmental and Economic Costs of Soil Erosion and Conservation Benefits. *Science* **1995**, *267*, 1117–1123. [CrossRef] [PubMed]
2. Thaler, E.A.; Larsen, I.J.; Yu, Q. The Extent of Soil Loss across the US Corn Belt. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e1922375118. [CrossRef] [PubMed]
3. Forster, D.L.; Bardos, C.P.; Southgate, D.D. Soil Erosion and Water Treatment Costs. *J. Soil Water Conserv.* **1987**, *42*, 349–352.
4. McDowell, R.; Sharpley, A.; Folmar, G. Phosphorus Export from an Agricultural Watershed: Linking Source and Transport Mechanisms. *J. Environ. Qual.* **2001**, *30*, 1587–1595. [CrossRef] [PubMed]
5. de Vente, J.; Poesen, J.; Arabkhedri, M.; Verstraeten, G. The Sediment Delivery Problem Revisited. *Prog. Phys. Geogr. Earth Environ.* **2007**, *31*, 155–178. [CrossRef]
6. Pandey, A.; Himanshu, S.K.; Mishra, S.K.; Singh, V.P. Physically Based Soil Erosion and Sediment Yield Models Revisited. *Catena* **2016**, *147*, 595–620. [CrossRef]
7. Gelder, B.; Sklenar, T.; James, D.; Herzmann, D.; Cruse, R.; Gesch, K.; Laflen, J. The Daily Erosion Project—Daily Estimates of Water Runoff, Soil Detachment, and Erosion. *Earth Surf. Process. Landf.* **2018**, *43*, 1105–1117. [CrossRef]
8. Flanagan, D.C.; Gilley, J.E.; Franti, T.G. Water Erosion Prediction Project (WEPP): Development History, Model Capabilities, and Future Enhancements. *Trans. ASABE* **2007**, *50*, 1603–1612. [CrossRef]
9. Tomer, M.D.; Porter, S.A.; James, D.E.; Boomer, K.M.B.; Kostel, J.A.; McLellan, E. Combining Precision Conservation Technologies into a Flexible Framework to Facilitate Agricultural Watershed Planning. *J. Soil Water Conserv.* **2013**, *68*, 113A–120A. [CrossRef]
10. Reybold, W.U.; TeSelle, G.W. Soil Geographic Data Bases. *J. Soil Water Conserv.* **1989**, *44*, 28–29.
11. 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. [CrossRef]
12. Laamrani, A.; Joosse, P.; Feisthauer, N. Determining the Number of Measurements Required to Estimate Crop Residue Cover by Different Methods. *J. Soil Water Conserv.* **2017**, *72*, 471–479. [CrossRef]
13. 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.
14. 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]
15. Dressing, S.; Harcum, J. *Recommendation Report for the Establishment of Uniform Evaluation Standards for Application of Roadside Transect Surveys to Identify and Inventory Agricultural Conservation Practices for the Chesapeake Bay Program Partnership's Watershed Model*; Chesapeake Bay Program: Annapolis, MD, USA, 2017.
16. Pilger, N.; Berg, A.; Joosse, P. Semi-Automated Roadside Image Data Collection for Characterization of Agricultural Land Management Practices. *Remote Sens.* **2020**, *12*, 2342. [CrossRef]
17. Laflen, J.M.; Amemiya, M.; Hintz, E.A. Measuring Crop Residue Cover. *J. Soil Water Conserv.* **1981**, *36*, 341–343.
18. Ding, Y.; Zhang, H.; Wang, Z.; Xie, Q.; Wang, Y.; Liu, L.; Hall, C.C. A Comparison of Estimating Crop Residue Cover from Sentinel-2 Data Using Empirical Regressions and Machine Learning Methods. *Remote Sens.* **2020**, *12*, 1470. [CrossRef]
19. Gelder, B.K.; Kaleita, A.L.; Cruse, R.M. Estimating Mean Field Residue Cover on Midwestern Soils Using Satellite Imagery. *Agron. J.* **2009**, *101*, 635–643. [CrossRef]
20. 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]

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. Van Deventer, A.P.; Ward, A.D.; Gowda, P.M.; Lyon, J.G. Using Thematic Mapper Data to Identify Contrasting Soil Plains and Tillage Practices. *Photogramm. Eng. Remote Sens.* **1997**, *63*, 87–93.

23. Zhang, Y.; Du, J. Improving Maize Residue Cover Estimation with the Combined Use of Optical and SAR Remote Sensing Images. *Int. Soil Water Conserv. Res.* **2023**, *12*, 578–588. [CrossRef]

24. Maxwell, A.E.; Warner, T.A.; Fang, F. Implementation of Machine-Learning Classification in Remote Sensing: An Applied Review. *Int. J. Remote Sens.* **2018**, *39*, 2784–2817. [CrossRef]

25. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [CrossRef]

26. Hively, W.D.; Duiker, S.; McCarty, G.; Prabhakara, K. Remote Sensing to Monitor Cover Crop Adoption in Southeastern Pennsylvania. *J. Soil Water Conserv.* **2015**, *70*, 340–352. [CrossRef]

27. Bull, L. Residue and Tillage Systems for Field Crops. Resources and Technology Division, Economic Research Service, U.S. Department of Agriculture: Washington, DC, USA, 1993.

28. Roy, D.P.; Kovalskyy, V.; Zhang, H.K.; Vermote, E.F.; Yan, L.; Kumar, S.S.; Egorov, A. Characterization of Landsat-7 to Landsat-8 Reflective Wavelength and Normalized Difference Vegetation Index Continuity. *Remote Sens. Environ.* **2016**, *185*, 57–70. [CrossRef]

29. Beeson, P.C.; Daughtry, C.S.T.; Wallander, S.A. Estimates of Conservation Tillage Practices Using Landsat Archive. *Remote Sens.* **2020**, *12*, 2665. [CrossRef]

30. Soil Survey Staff USDA NRCS Web Soil Survey. Available online: https://websoilsurvey.sc.egov.usda.gov (accessed on 31 May 2021).

31. 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]

32. Zylman, J.; Weindorf, D.C.; Wittie, R.; McFarland, A.; Butler, T. Field-Truthing of USDA-Natural Resources Conservation Service Soil Survey Geographic Data on Hunewell Ranch, Erath County, Texas. *Soil Surv. Horiz.* **2005**, *46*, 135–145. [CrossRef]

33. Laamrani, A.; Joosse, P.; McNairn, H.; Berg, A.A.; Hagerman, J.; Powell, K.; Berry, M. Assessing Soil Cover Levels during the Non-Growing Season Using Multitemporal Satellite Imagery and Spectral Unmixing Techniques. *Remote Sens.* **2020**, *12*, 1397. [CrossRef]

34. Upadhyay, P.C.; Lory, J.A.; DeSouza, G.N.; Lagaunne, T.A.; Spinka, C.M. Classification of Crop Residue Cover in High-Resolution RGB Images Using Machine Learning. *J. ASABE* **2022**, *65*, 75–86. [CrossRef]

35. Daughtry, C.S.T.; Hunt, E.R. 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]

36. Gowda, P.H.; Dalzell, B.J.; Mulla, D.J.; Kollman, F. Mapping Tillage Practices with Landstat Thematic Mapper Based Logistic Regression Models. *J. Soil Water Conserv.* **2001**, *56*, 91–96.

37. Hively, W.D.; Shermeyer, J.; Lamb, B.T.; Daughtry, C.T.; Quemada, M.; Keppler, J. Mapping Crop Residue by Combining Landsat and Worldview-3 Satellite Imagery. *Remote Sens.* **2019**, *11*, 1857. [CrossRef]