Use of high resolution unmanned aerial systems imagery and machine learning to evaluate grain sorghum tolerance to mesotrione

Isaac Barnhart 1, Sushila Chaudhari 2, Balaji Aravindhan Pandian 1, Pagadala Venkata Vara Prasad 3, Ignacio A. Ciampitti 1, and Mithila Jugulam 1,*

1 Department of Agronomy, Kansas State University, Manhattan, Kansas, United States of America; ihb@ksu.edu (I.B.); aravindhan@ksu.edu (B.A.P); ciampitti@ksu.edu (I.A.C); mithila@ksu.edu (M.J.)
2 Department of Horticulture, Michigan State University, East Lansing, Michigan, United States of America; sushilac@msu.edu
3 Sustainable Intensification Innovation Lab, Department of Agronomy, Kansas State University, Manhattan, Kansas, United States of America; vara@ksu.edu (P.V.V.P.)
* Correspondence: mithila@ksu.edu; Tel.: 1-785-532-2755

Abstract: Manual evaluation of crop injury to herbicides is time-consuming. Unmanned aircraft systems (UAS) and high-resolution multispectral sensors and machine learning classification techniques have the potential to save time and improve precision in the evaluation of herbicide injury in crops, including grain sorghum (Sorghum bicolor L. Moench). The objectives of this research were to (1) evaluate three supervised classification algorithms (support vector machine, maximum likelihood, and random forest) for categorizing high-resolution UAS imagery to aid in data extraction and (2) evaluate the use of vegetative indices (VIs) collected from UAS imagery as an alternative to traditional methods of visual herbicide injury assessment in mesotrione-tolerant grain sorghum breeding trials. An experiment was conducted in a randomized complete block design using a factorial treatment arrangement of three genotypes by four mesotrione doses. Herbicide injury was rated visually on a scale of 0 (no injury) to 100 (complete plant mortality). The UAS flights were flown at 9, 15, 21, 27, and 35 days after treatment. Results show the SVM algorithm to be the most consistently accurate, and high correlations (r = -0.83 to -0.94; p < 0.0001) were observed between the normalized difference vegetative index (NDVI) and ground-measured herbicide injury. Therefore we conclude that VIs collected with UAS coupled with machine learning image classification, has the potential to be an effective method of evaluating mesotrione injury in grain sorghum.

Keywords: unmanned aerial vehicle, grain sorghum, herbicide injury, remote sensing, sorghum breeding

1. Introduction

Grain sorghum (Sorghum bicolor L. Moench) is an important crop with diverse uses throughout the world [1]. Sorghum is known to have originated from sub-
Saharan Africa, where it has been used as a major food crop [2]. Grain sorghum has a vast genetic diversity with many genotypes possessing agronomically-desirable traits [3]. Traditionally, sorghum is used in the Western Hemisphere as animal feed; however, other uses include biofuel, forage production, and as a gluten-free alternative for human consumption [3-4]. In the United States, grain sorghum ranks the fifth-most important grain crop with 1.9 million hectares harvested in 2019 [5]. As sorghum is adapted to semi-arid regions of the world, it has been shown to have distinct yield advantages over corn (Zea mays L.) in such environments [6].

Weed competition is a common biotic pestilence that interferes with grain sorghum production. Prior research has demonstrated that infestations of weeds in grain sorghum can reduce yields ranging from 8 to 56%, depending on the type of weeds [7]. Despite other methods of weed control such as crop rotation and row cultivation, herbicides are the most common method of weed control in grain sorghum in the United States [7-8]. Several herbicides are registered for use in grain sorghum as pre (PRE)- or post (POST)-emergence treatments. For example, herbicides such as atrazine and mesotrione can be used as a PRE, and 2,4-D and dicamba as POST treatments [9]. Despite having several herbicides used for PRE treatments, the options for POST treatments are limited in grain sorghum, especially those used for grass weed control. More options for POST grass control herbicides are needed for grain sorghum cropping systems.

The 4-hydroxyphenylpyruvate dioxygenase (HPPD)-inhibiting herbicides (e.g. mesotrione, tembotrione) are widely used for POST emergence grass weed control in corn [10] but not registered for POST use in grain sorghum because of crop injury [11]. These herbicides are used both as soil and foliar applied, and control a broad spectrum of grass and broadleaf species while providing soil residual activity for extended protection [10, 12-13]. These herbicides inhibit the HPPD enzyme in the plastoquinone biosynthesis pathway, leading to the depletion of plastoquinone levels. This results in the inhibition of carotenoids biosynthesis and subsequent plant death by photo-oxidation of chloroplasts. Because of this photo-oxidation, the main symptom after treatment with these herbicides is the bleaching of plant tissue, although additional symptoms including stunting of growth, leaf chlorosis and necrosis are also common in susceptible plants [8, 10]. HPPD-inhibiting herbicides are widely used as POST in corn, which can actively metabolize these herbicides into non-phytotoxic compounds [10].

Mesotrione is an HPPD-inhibitor in the triketoone chemical family [14], and can be used as a pre-emergence treatment in grain sorghum. However, it has been shown to cause damage including 20% chlorosis in sorghum when applied POST [8, 15]. Recent research has identified grain sorghum genotypes with elevated tolerance to post-emergence applications of mesotrione with minimal crop damage [16]. These grain sorghum genotypes are valuable for the development of mesotrione-tolerant varieties in breeding programs. Although herbicide-tolerant traits could be quite useful for growers, implementing herbicide tolerance to grain sorghum breeding operations can be difficult. This is because several combinations
of herbicides rates and genetic lines need to be fully investigated to quantify the data of herbicide injury level [17]. In addition, previous methods to assess plant responses to herbicides, including visual scoring, portable chlorophyll sensors, and biomonitoring [18-20] are labor-intensive and time-consuming. New methods of evaluation of herbicide damage are warranted to support and hasten the screening large number of genotypes and populations in the breeding programs for the development of herbicide-tolerant technology in grain sorghum.

The use of high-resolution remote sensing techniques coupled with machine learning image classification algorithms is one of the promising methods for quantification of plant response to various biotic or abiotic stresses [21-22]. Image classification via remote sensing has been explored extensively, with the majority of platforms using either satellite, piloted aircraft, or unmanned aerial systems (UAS) [19, 23]. Images collected from these platforms can then be separated into distinct classes through various algorithms, allowing extraction of data from individual classes for external analysis [24-26]. A highly accurate sect of image classification includes a supervised approach, which uses user-defined training samples to make these feature classifications [23, 27]. In addition to image classification, vegetative indices such as the normalized difference vegetation index (NDVI) can be computed, which have been shown to be useful in predicting plant parameters such as biomass, nitrogen status and chlorophyll content [28-30]. The quality of the data enhances with improved image resolution, making UAS a useful platform for collecting high-resolution remote sensing imagery in agricultural settings [31-34].

Proposed uses of UAS imagery in weed science, include weed pressure mapping, quantifying herbicide damage on non-target crops, herbicide applications, and site-specific weed management [32-33, 35-38]. However, very few studies have been carried out using UAS imagery to quantify herbicide tolerance in grain sorghum. Therefore, the objectives of this study were to (1) specify a methodology to aid in the image processing, data extraction, and statistical analysis of mesotrione tolerant herbicide trials, including evaluating three algorithms for image classification, and (2) investigate the use of UAS imagery as an alternative to traditional methods of assessing visual mesotrione injury in grain sorghum.

2. Materials and Methods

2.1. Field Experiment

A field study was conducted in 2019 at the Kansas State University Ashland Bottoms Research Farm in Manhattan, Kansas (Figure 1). The soil type at the study location was a Reading silt loam with 2.42% organic matter and a pH of 6.07. A grain sorghum genotype (G-1) identified with tolerance to mesotrione [16] was planted along with a known mesotrione-susceptible genotype (S-1) and a Pioneer® (Corteva Agriscience, Wilmington, Delaware; USA) sorghum hybrid (84G62) for comparison. The experiment was arranged in a randomized complete block design (with three
replications) with a two-way factorial arrangement of treatments, consisting of three genotypes and four herbicide rates. The genotypes were planted on June 8, 2019 at a rate of 173,000 seeds ha$^{-1}$, with a row spacing of 76 cm and a planting depth of 5 cm. The experiment plots were 2 m wide and 6.5 m long, consisting of four sorghum rows. The two middle rows consisted of each genotype (G-1, S-1, 84G62), and two outside border rows of Pioneer 84G62 were planted to act as a buffer between treatments. Mesotrione (Callisto® SC; Syngenta Crop Protection, LLC, Greensboro, North Carolina, USA) rates of 0, 105, 420, and 840 g ai ha$^{-1}$ were applied at the 3-4 leaf growth stage using a CO$_2$-pressurized backpack sprayer consisting of a four nozzles boom fitted with 110-02 flat-fan nozzles (TeeJet Spraying Systems Co., Wheaton, Illinois, USA) calibrated to deliver 187 L ha$^{-1}$ at 276 kPa. Fertilizer applications followed the recommendations determined from a pre-plant soil test, and pests and diseases were controlled on an as-needed basis.

![Figure 1](image.jpg)

**Figure 1.** Location of the field trial in Manhattan, Kansas (US) evaluating 84G62, S-1, and G-2 spectral responses to mesotrione. The genotypes were planted in two rows, as denoted by the red rectangles. Plot numbers are denoted below each rectangle.

2.2. Image Acquisition

Prior to seedling emergence, 10 ground control points (GCPs) were placed within the experiment, and real-time kinematic (RTK) points were collected to aid
in image processing. In addition to GCPs, each plot received two diagonally-placed ground targets indicating the sub-plots where ground-truth measurements were taken. Dimensions of the sub-plots were approximately 1.5 m x 1m. After processing, geo-centered border polygons could then be drawn over these sub-plots to allow for data extraction (Figure 2).

The Imagery was collected with a DJI Matrice 200 (DJI Inc., Shenzhen, China; https://www.micasense.com/) multi-rotor aircraft equipped with a Micasense RedEdge-MX multispectral camera (Micasense Inc., Seattle, Washington, USA; https://www.micasense.com/). The camera is capable of capturing five spectral bands of the visible and invisible electromagnetic spectrum (blue, 465-485 nm; green, 550-570 nm; red, 663-673 nm; red edge, 712-722 nm; near infrared, 820-860 nm [40]). Each band is captured independently with a separate camera lens, and is capable of capturing images with a spatial resolution of 8 cm/pixel at an altitude of 120 m. To measure fluctuations in ambient lighting conditions, a down-welling light sensor (DLS) was included on the top of the aircraft.

Flights were flown 9, 15, 21, 27, and 35 days after treatment (DAT) for data collection. Flights were flown under clear, sunny conditions with no cloud cover. These conditions were selected to standardize lighting across all measurement dates, ensuring differences in data due to lighting conditions would be minimized. Flight dates were taken in order to obtain data as close to weekly measurements as
possible, as dictated by our chosen lighting conditions. All flights were taken within ± 2.5 hours of solar noon as recommended by the camera manufacturers. Radiometric calibration was performed using the calibration panel before and after each flight to ensure image quality. Each flight was flown at an altitude of 15 m and was flown using a pre-programmed GPS waypoint mapping mission using the DJI Pilot application. To ensure image quality, each flight was flown with an 80% front overlap and a 75% side overlap. The camera was set to capture images every 2 seconds to ensure adequate amounts of images were taken for processing. Images collected in flight were stored on an SD card and transferred to a desktop computer for further processing.

2.3. Image Processing

Data processing followed a workflow involving image orthomosaic generation in Agisoft Metashape (Agisoft LLC., St. Petersburg, Russia) and data extraction in ArcGIS Pro (ESRI Headquarters, Redlands, California, USA) (Figure 3). UAS images processed using Metashape were processed with the slight modifications discussed in Holman et al. [39]: When aligning images and generating the sparse point cloud, alignment accuracy was set to high as opposed to ultra-high to reduce processing time and decrease levels of noise. When generating the dense point cloud, depth filtering was disabled to prevent any smoothing of the sorghum canopy. In addition, the quality of the dense point cloud was also set to high instead of ultra-high to reduce processing time. The resulting orthomosaic spatial resolution was between 1.1-1.3 cm/pixel for each of the 5 flight dates.
Remote Sens. 2020, 12, x FOR PEER REVIEW

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 October 2020
doi:10.20944/preprints202010.0022.v1

Figure 3. Image processing and data extraction workflow.

The image orthomosaic was exported to ArcGIS Pro for further data analysis. Ground targets within each plot were located, and sub-plots were defined. The sub-plots were then clipped from the original image, and vegetative indices (VIs) were computed on the sub-plots. The vegetative indices (VIs) computed were NDVI, enhanced normalized difference vegetation index (ENDVI), simple ratio (SR), and enhanced vegetation index 2 (EVI2) (Table 1). These VIs were chosen for their close relations to plant greenness and canopy chlorophyll content [41].

Table 1. Equations and common uses for vegetative indices used in this study.

<table>
<thead>
<tr>
<th>Vegetative Index</th>
<th>Formula</th>
<th>Use</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDVI</td>
<td>$\frac{\rho_{\text{NIR}} - \rho_R}{\rho_{\text{NIR}} + \rho_R}$</td>
<td>Green biomass, chlorophyll</td>
<td>[40, 42-44]</td>
</tr>
<tr>
<td>ENDVI</td>
<td>$\frac{((\rho_{\text{NIR}} + \rho_G) - (2\times \rho_R))}{((\rho_{\text{NIR}} + \rho_G) + (2\times \rho_R))}$</td>
<td>Chlorophyll content</td>
<td>[41, 44]</td>
</tr>
<tr>
<td>SR</td>
<td>$\frac{\rho_{\text{NIR}}}{\rho_R}$</td>
<td>Chlorophyll, leaf area index</td>
<td>[45]</td>
</tr>
<tr>
<td>EVI2</td>
<td>$2.5 \times \frac{\rho_{\text{NIR}} - \rho_R}{\rho_{\text{NIR}} + (2.4 \times \rho_R) + 1}$</td>
<td>Chlorophyll, canopy cover</td>
<td>[41, 43, 46-47]</td>
</tr>
</tbody>
</table>

Acronyms: NDVI, normalized difference vegetation index; ENDVI, enhanced normalized difference vegetation index; SR, simple ratio; EVI2, enhanced vegetation index 2. $\rho$ indicates spectral reflectance.
2.4. Machine Learning Classification

A supervised classification approach was chosen over unsupervised methods because of the classification accuracy advantage over unsupervised methods [26-27]. Due to high image resolutions and standardized lighting conditions on each measurement date, we chose to use pixel-based classification algorithms, as opposed to other methods such as object based. After computing VI imagery, each sub-plot was classified into three categories: Sorghum leaves, shadows, and soil. To test for differences between accuracies in classification algorithms, we chose three different algorithms: support vector machine (SVM), random forest (RF), and maximum likelihood (ML). These three algorithms were chosen because they are all the options offered for pixel-based classifications in ArcGIS Pro. In order to classify the imagery, we followed a similar approach described by Tay et al. and Makanza et al. [47-48]. First, a classification scheme consisting of three categories was created. Twenty representative training samples were taken at random from the sub-plot, grouping similar pixels based on visual similarity and assigning them to each class. The training dataset and classification schema was then saved as a signature file and used to classify each sub-plot using each machine learning algorithm. In order to achieve maximum accuracy across each measurement day, this process was repeated with new representative training samples generated for each day. All algorithms were run with the default settings provided by ArcGIS Pro.

2.5. Ground Truth Measurements

Herbicide injury was assessed visually by using a scale of 0 (no visual injury) to 100 (complete plant mortality). Injury ratings were based on the presence of foliar bleaching, growth stunting, leaf chlorosis and tissue necrosis [8, 10, 49]. Ratings were taken within sub-plots so data from flights could be taken from the same plants in which the ratings were assigned. Each rating was taken ± 1 day of flight measurements.

Accuracy assessments were conducted on the classification methods by computing confusion matrices [50-51]. For each measurement day, 300 accuracy assessment points were generated via stratified random sampling. Each point was manually designated to the class to which it belonged, relative to the original orthomosaic image, creating ground reference points. When overlaid onto each classified raster, overall accuracy (OA) percentages were then computed by dividing the total number of correctly classified pixels by the total number of reference pixels. Accuracies at or above 85% were considered to be accurate classifications, which is a commonly-used target accuracy when classifying imagery [50]. For this study, we selected the most consistently-accurate algorithm across all 5 treatment dates to use for further data extraction (see results section).
2.6. Data Extraction

The process of data extraction from each sub-plot is shown in Figure 4. To prevent extraction from background features, data from each sub-plot was extracted using a conditional statement. Using the selected classified image, a conditional statement was built to allow for data from each computed VI to be extracted only from the sorghum plants. The average value of each VI was then extracted and exported for further analysis.

Figure 4. Methodology for vegetative indices (VI) data extraction based on the results of the image classification. (A) Original red-green-blue (RGB) sub-plot image; (B) each VI was computed from the original RGB images, creating a new raster layer for each VI; (C) RGB images were classified via supervised, pixel-based algorithms into leaves, shadows, and soil; and (D) each VI raster was combined with the classified image to create a new raster layer, allowing for VI data from only the sorghum leaves to be extracted for statistical analysis.

2.7. Statistical Analysis

All data were exported for analysis in R statistical program (R Core Team, Vienna, Austria; https://www.R-project.org). Data analysis was completed in two stages. First, statistical differences between mean overall classification accuracies (per algorithm) were tested with the Shapiro-Wilks test to verify assumptions of analysis of variance (ANOVA) [52]. The classification accuracy dataset failed to meet ANOVA assumptions and was therefore analyzed using the Kruskal-Wallis chi-squared goodness of fit test [53]. Significant differences between algorithms were found with a post-hoc analysis using the Dunn test [54]. Second, relationships between VIs and ground truth injury ratings were investigated and tested for
significance with Pearson correlation coefficients [40, 48, 55-57]. Correlations were then used to detect differences in spectral values between genotypes and application rates. To examine statistical difference in spectral responses of genotypes, a two-way ANOVA was conducted comparing the effects of genotype/hybrid and rate on VI values, with means being separated using a Tukey Honest Significant Difference test. All significance levels were set at $\alpha = 0.05$.

3. Results

3.1. Classification Accuracy

Overall classification accuracies ranged from 82 to 93% across algorithms, which can be seen in Table 2. The SVM was the most consistent across all dates with accuracies of 90% (27 DAT) to 92% (9 DAT). The algorithm with the most variability was ML, while the lowest accuracies were achieved with RF (82%, 15 DAT; 88%, 9 DAT). The Kruskal-Wallis test revealed significant differences between groups (Chi-square = 7.41, p-value = 0.025), and the RF algorithm was found to be less accurate when compared with SVM and ML (Table 3). Because SVM was consistently the most accurate classification algorithm, it was chosen to continue to extract data for further analysis. When comparing the overall accuracy means of each algorithm, all algorithms accurately classified the imagery into three established classes (leaves, shadows, and soil), as compared to the threshold of 85%. However, RF was found to be significantly less accurate than SVM and ML, suggesting that RF may not be the best choice in similar supervised, pixel-based classifications.

<table>
<thead>
<tr>
<th>DAT</th>
<th>SVM</th>
<th>RF</th>
<th>ML</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>92</td>
<td>88</td>
<td>85</td>
</tr>
<tr>
<td>15</td>
<td>91</td>
<td>82</td>
<td>92</td>
</tr>
<tr>
<td>21</td>
<td>91</td>
<td>87</td>
<td>93</td>
</tr>
<tr>
<td>27</td>
<td>90</td>
<td>87</td>
<td>92</td>
</tr>
<tr>
<td>35</td>
<td>91</td>
<td>82</td>
<td>90</td>
</tr>
</tbody>
</table>

Table 2. Overall accuracies for each algorithm on each measurement date.

Acronyms: DAT, days after treatment; SVM, support vector machine; RF, random forest; ML, maximum likelihood. Accuracies are expressed in percentages.
Table 3. Results of the Dunn test indicating significant differences between mean overall classification accuracy percentages. No significant differences were found between the SVM-ML algorithms, but the RF algorithm was significantly less than the others.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>P-value</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>ML-RF</td>
<td>0.02</td>
<td>*</td>
</tr>
<tr>
<td>ML-SVM</td>
<td>0.94</td>
<td>ns</td>
</tr>
<tr>
<td>RF-SVM</td>
<td>0.02</td>
<td>*</td>
</tr>
</tbody>
</table>

ML, maximum likelihood; RF, random forest; SVM, support vector machine.

1 Asterisks indicate significant differences (p < 0.05); “ns” indicates non-significance.

3.2. Relation of Vegetative Indices to Ground Truth Injury

Significant negative correlations were observed across all measurement dates with each VI. High negative correlations were observed between UAS-based VI data and ground truth herbicide injury across all measurement dates (Table 4). The NDVI coefficients (r = -0.94 to -0.83) consistently displayed the highest correlations with visual injury symptoms on every measurement day. ENDVI (r = -0.92 to -0.82) and EVI2 (r = -0.94 to -0.82) correlations were very similar to NDVI values and thus were also very highly correlated with injury symptoms. The SR demonstrated the most variability in each measurement date (r = -0.92 to -0.70). In addition to VI data, individual spectral bands were tested for significance with ground-measured injury scores, with significances not higher than what was determined for VIs (data not shown).

Table 4. Pearson correlation coefficients between vegetative indices and ground truth visual injury ratings.

<table>
<thead>
<tr>
<th>VI</th>
<th>9 DAT</th>
<th>15 DAT</th>
<th>21 DAT</th>
<th>27 DAT</th>
<th>35 DAT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r</td>
<td>p-value</td>
<td>r</td>
<td>p-value</td>
<td>r</td>
</tr>
<tr>
<td>NDVI</td>
<td>-0.94</td>
<td>&lt; 0.0001</td>
<td>-0.91</td>
<td>&lt; 0.0001</td>
<td>-0.85</td>
</tr>
<tr>
<td>ENDVI</td>
<td>-0.92</td>
<td>&lt; 0.0001</td>
<td>-0.90</td>
<td>&lt; 0.0001</td>
<td>-0.84</td>
</tr>
<tr>
<td>EVI2</td>
<td>-0.94</td>
<td>&lt; 0.0001</td>
<td>-0.90</td>
<td>&lt; 0.0001</td>
<td>-0.84</td>
</tr>
<tr>
<td>SR</td>
<td>-0.92</td>
<td>&lt; 0.0001</td>
<td>-0.83</td>
<td>&lt; 0.0001</td>
<td>-0.79</td>
</tr>
</tbody>
</table>

As NDVI was shown to be the most consistently related to mesotrione injury, it was chosen for the ANOVA analysis. Across all treatment dates, there was no significant interaction effect between genotype and rate (data not shown). Despite this, the main effect models demonstrated significant differences, with the results of the post-hoc test shown in Figure 5. When looking at the main effects model for genotypes, the NDVI response for the G-1 genotype across all mesotrione doses was not statistically different from the commercial hybrid response until 35 DAT. However, in all cases, the NDVI values for the G-1 genotype and commercial hybrid
were statistically different from the S-1 genotype, which was to be expected given the S-1 susceptibility to mesotrione. For the rate main effect model, plants treated with 0 and 105 g ae ha⁻¹ were not statistically different from one another on each measurement date, indicating that a low dose of mesotrione did not significantly injury the lines used in this study. Interestingly, on the 27 and 35 DAT, it appeared that the plots treated with a higher rate of mesotrione (420 and 840 g ai ha⁻¹) started to recover from their injuries, but were always significantly lower than the control and 105 g ae ha⁻¹ dose.
Figure 5. ANOVA results for main effect models, genotype and rate. Means were separated using Tukey’s HSD test at the 0.05 significance level, and error bars denote standard error of the mean.
4. Discussion

UAS imagery has been used to monitor various agronomic traits, including plant response to stress [48]. Supervised machine learning algorithms are commonly used to classify remote sensing imagery, and have been shown to be highly accurate in terms of overall classification accuracies [24, 58]. Our current research shows that machine learning algorithms, most notably the SVM algorithm, could be an effective method of extracting VI values from grain sorghum treated with different rates of mesotrione (Table 3).

Key to this finding is the use of high-resolution UAS imagery coupled with plots across uniform lighting conditions. Within each plot, there were three classes identified in each classification schema: leaves, shadows, and soil. A stark contrast between pixel brightness values occurred throughout each of the classes, allowing for the SVM algorithm to effectively segregate each class throughout the plots with high accuracy percentages. As SVM and ML were not statistically different in terms of overall accuracy, relationships between VI data and ground-measured injury are not expected to be statistically different if ML were to be used instead of SVM for data extraction. We expect that this classification schema could be used in grain sorghum to assess injury of other HPPD-inhibitor herbicides, such as tembotrione, because similar classes would be expected, regardless of the location of study or genotypes. As pixel-based classifications were studied, we did not compare these results with other methods of segregating vegetation from background noise, including using the Otsu algorithm [59] for binary thresholding and object-based image analysis. Further research for vegetation classification in screening for herbicide tolerance studies should compare classification accuracies of pixel-based, object-based, and binary thresholding algorithms.

The classification method chosen for this study was supervised, which are often more accurate than unsupervised in some situations [27]. One of the main benefits of unsupervised classifications is that large areas of land can be classified in a very short time frame [60]. In addition, as supervised classification methods require a training dataset to be collected from the imagery itself, unsupervised classification algorithms would be able to bypass this step and potentially hasten the classification process [61]. However, as demonstrated by our study, segregating sorghum vegetation in herbicide breeding trials does not require many classes or training samples due to the relative homogeneity of features within the field. 20 training samples were used here, as is the recommended minimum sample size for parametric classifiers such as ML [62]. However, it is entirely possible that an even smaller amount of samples could be used when using SVM, which is much less sensitive to background interference and can be used when few training samples are present [63]. As supervised classification algorithms performed well in this study, more research is needed to compare supervised classifications with unsupervised classifications in herbicide tolerance trials to see if statistically significant differences exist between overall accuracies.
Using the NDVI values and an ANOVA, it was observed that differences in spectral values existed for genotypes with respect to mesotrione application rates. The ability to detect spectral responses of plots treated with mesotrione could aid in quickly assessing the variable responses of multiple sorghum genotypes to mesotrione treatment. To be useful for sorghum breeders, differences in plant responses to mesotrione must be detected by the camera and able to be seen after data extraction. Even though the EVI2 and ENDVI indices were found to be similar in correlations with NDVI, we chose to evaluate only the NDVI as it is one of the most widely used vegetative indices [64]. This study could be improved upon by testing other indices in future studies. Our data demonstrates that a multispectral camera was able to detect differences among genotype responses, which could allow breeders to know which genotypes are responding to herbicide treatment. This prevents a large-scale assessment of the entire operation, leading to reduced time, labor, and resources to achieve the desired objective.

It should be noted that all VIs showed the strongest correlations on 9 and 15 DAT, and showed slightly weaker correlations for the rest of the measurement days. This is most likely because reductions in chlorophyll due to herbicide injury were clearly visible on the canopy early during the experiment, but were overshadowed by new growth as the growing season progressed. As a result, injury symptoms were still visible during ground truth ratings but were not visible to the camera. As healthy plants absorb light in the visible spectrum and reflect near infrared (NIR) energy, higher NDVI values are observed with healthier plants, whereas lower NDVI values are recorded as plant health declines [36]. This explains the negative correlations, as plants injured by mesotrione applications experienced a loss of green pigmentation [65]. In addition, bleaching symptoms due to mesotrione were not directly classified with machine learning algorithms due to spectral similarities with soil values. Instead, the reduction in leaf greenness showed to be sufficient in determining correlations with ground truth injury ratings.

VIs have been successfully correlated with herbicide injury involving chlorophyll pigmentation loss in previous research. Huang et al. [66] found that NDVI values were highly related to cotton (Gossypium hirsutum L.) injury from aerial applications of glyphosate, with very strong correlations observed between increased frequencies of spray drift droplets and NDVI values. Dicke et al. [35] found reductions in NDVI values in plots treated with sulfonylurea herbicides as well as significant correlations between NDVI values and corn yield, suggesting that low NDVI values in herbicide-damaged plots can be used as predictor of final yield. Duddu et al. [67] reported strong correlations between the optimized soil adjusted vegetation index (OSAVI) and visual injury ratings of faba beans (Vicia faba L.) treated with 9 different herbicide tank mixtures. Faba bean visual injury ratings were based on visible growth reduction and tissue chlorosis. Our data is consistent with these findings in that the increase in herbicide foliar injury is strongly related to changes in VI scores. As a final recommendation for future studies, relationships between UAS data and other herbicide modes of action
should be determined to investigate how useful UAS would be for determining grain sorghum tolerance to other herbicides.

Herbicide injury in grain sorghum breeding trials is traditionally been assessed with visual injury ratings, which can be prone to error due to variations among evaluators. For trials consisting of thousands of grain sorghum plots, traditional injury assessment methods would take many hours or days to complete. Alternatively, UAS imagery can be taken in as little as 20-25 minutes, and while the data extraction methodology may seem to take a long time, the methodology can actually be completed quickly once the process is established.

Limitations of this study that could be addressed in future studies are centered on the lack of use of more number of herbicides, genotypes tested and geographic restrictions. As there was only one herbicide used in this study, it remains to be determined if tolerance/susceptibility to other herbicides could be detected with UAS. Mesotrione injury was well-detected due to foliar bleaching symptoms, resulting from the destruction of phyto-oxidation of chloroplasts. This resulted in stark differences between treated and untreated plants, which were able to be detected with the camera due to differences in chlorophyll pigmentation loss. As other herbicides do not directly result in stark pigmentation losses (i.e. auxins), the ability to detect sorghum responses to these herbicides remains to be determined.

Additional limitations involve use of genotypes that are not bred similarly for cultivation. While Pioneer 84G62 is a commercially grown hybrid with desirable agronomic traits, the genotypes G-1 and S-1 represent sorghum diversity panel which are do not possess agronomically suitable traits. The other limitation includes, use of only one year data at one geographic location. This experiment was conducted in one field (Manhattan, KS), and has only been replicated across one year. This methodology remains to be tested across multiple years/locations to determine the robustness of detecting differences in mesotrione tolerance. Further studies should determine UAS injury characterization across multiple locations and growing conditions.

Although significant spectral difference existed between certain genotypes and rates of herbicide applications, the differences were not as pronounced as would be expected between mesotrione tolerant and susceptible genotypes. A likely explanation can be found in precipitation patterns. Rainfall was plentiful for the 2019 growing season in central KS, enabling adequate growth and development for all sorghum lines studied. This further demonstrates the need for additional replications across different geographic regions, as greater spectral differences between tolerant and susceptible genotypes could potentially be observed in different growing conditions.
5. Conclusions

This study suggests a methodology to aid in the image processing, data extraction, and statistical analysis for future research with multispectral evaluation of herbicide tolerance in grain sorghum. Additionally, we suggest that VIs (notably NDVI in this study) collected from high-resolution imagery, coupled with image classification, may be a useful tool for sorghum breeders to quantify the degree of mesotrione injury in grain sorghum. Across all measurement dates, a supervised, pixel-based SVM classifier was shown to be accurate and consistent when classifying imagery into sorghum leaves, shadows, and soil. The high correlation between VI data and ground-measured injury scores indicates that high-resolution multispectral imagery can detect reductions in chlorophyll due to mesotrione injury. As NDVI values were shown to be the most correlated to herbicide injury, we have demonstrated that the most widely-used VI in agriculture is capable of detecting these changes. Therefore, cost-effective NDVI sensors of varying types could potentially be used to measure mesotrione injury or possibly other herbicide injury symptoms that can be quantifiable. Further research is necessary to determine the ability of different sensors to measure such injury, as well as compare other means of vegetation segmentation such as using the Otsu algorithm or object-based image analysis. Additional research is also necessary to determine if high-resolution imagery is capable of detecting damage symptoms of herbicides with differing modes of action.

Author Contributions: Conceptualization, M.J, S.C., B.A.P, I.B.; methodology, I.B., S.C.; validation, M.J., P.V.V.P, I.A.C.; formal analysis, I.B.; investigation, M.J.; writing – original draft preparation, I.B.; writing – review and editing, M.J., S.C., B.A.P., P.V.V.P., I.A.C.; visualization, I.B.; supervision, M.J., P.V.V.P., I.A.C.; project administration, M.J.; funding acquisition, M.J., I.A.C. All authors have read and agreed to the published version of the manuscript.

Funding: Research assistantship for I.B and B.A.P from the Center for Sorghum Improvement and the Department of Agronomy, Kansas State University is appreciated.

Acknowledgments: This is contribution number 20-329-J from the Kansas Agricultural Experiment Station, Kansas State University, Manhattan, Kansas.

Conflicts of Interest: The authors declare no conflict of interest.

References


66. Huang, Y.; Thomson, S.J.; Ortiz, B.V.; Reddy, K.N.; Ding, W.; Zablotsowicz, R.M.; Bright, J.R. Airborne remote sensing assessment of the damage to cotton caused by spray drift from aerially applied glyphosate through spray deposition measurements. *Biosyst Eng.* 2010, 107, 212-220. doi: 10.1016/j.biosystemseng.2010.08.003