Precision Weed Management Based on UAS Image Streams, Machine Learning, and PWM
Sprayers

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Geosciences

by

Jason A. Davis
University of Arkansas
Bachelor of Science in Agriculture Education, 2009
University of Arkansas
Master of Science in Agriculture and Extension Education, 2011

December 2022
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

_____
Jason A. Tullis, Ph.D.
Dissertation Director

_____   _____
Michael B. Daniels, Ph.D.   Richard G. Ham, Ed.D.
Committee Member   Committee Member

_____
Terry N. Spurlock, Ph.D.
Committee Member

ABSTRACT

Weed populations in agricultural production fields are often scattered and unevenly distributed; however, herbicides are broadcast across fields evenly. Although effective, in the case of post-emergent herbicides, exceedingly more pesticides are used than necessary. A novel weed detection and control workflow was evaluated targeting *Palmer amaranth* in soybean (*Glycine max)* fields. High spatial resolution (0.4 cm) unmanned aircraft system (UAS) image streams were collected, annotated, and used to train 16 object detection convolutional neural networks (CNNs; RetinaNet, Faster R-CNN, Single Shot Detector, and YOLO v3) each trained on imagery with 0.4, 0.6, 0.8, and 1.2 cm spatial resolutions. Models were evaluated on imagery from four production fields containing approximately 7,800 weeds. The highest performing model was Faster R-CNN trained on 0.4 cm imagery (precision = 0.86, recall = 0.98, and F1-score = 0.91). A site-specific workflow leveraging the highest performing trained CNN models was evaluated in replicated field trials. Weed control (%) was compared between a broadcast treatment and the proposed site-specific workflow which was applied using a pulse-width modulated (PWM) sprayer. Results indicate no statistical ($p < .05$) difference in weed control measured one ($M = 96.22\%$, $SD = 3.90$ and $M = 90.10\%$, $SD = 9.96$), two ($M = 95.15\%$, $SD = 5.34$ and $M = 89.64\%$, $SD = 8.58$), and three weeks (M = 88.55, SD = 11.07 and M = 81.78%, SD = 13.05) after application between broadcast and site-specific treatments, respectively. Furthermore, there was a significant ($p < 0.05$) 48% mean reduction in applied area (m^2) between broadcast and site-specific treatments across both years. Equivalent post application efficacy can be achieved with significant reductions in herbicides if weeds are targeted through site-specific applications. Site-specific weed maps can be generated and executed using accessible technologies like UAS, open-source CNNs, and PWM sprayers.

ACKNOWLEDGEMENTS

DEDICATION

This dissertation is dedicated to my wife Amanda. Thank you for your continued love, support, and understanding in my endeavors. I love you more than I can express and am so blessed to share my life with such a Godly woman. To my children Grayson, Reagan, and Henry, I love each of you and your individual developing interests, talents, and personalities. I look forward to watching each of you tackle your own life's adventures and hope that my pursuits serve as a form of inspiration and encouragement. To my parents Allen and Karen who have always provided Godly counsel and encouragement in all my adventures. Thank you and I love you both.

TABLE OF CONTENTS

1.0   INTRODUCTION

Herbicides protect today's crops from weed competition, that when left uncontrolled,

reduce crop yield and profits. Weed populations in production fields can often be described as

scattered and unevenly distributed; however, herbicides are broadcast across fields evenly.

Although effective, in the case of post-emergent herbicides, this method often uses exceedingly

more pesticides than necessary. Alternatively, applying pesticide only to locations that contain

weeds would also be effective, more cost-efficient, and more environmentally responsible

compared to broadcasting.

Modern agriculture equipment can place precise amounts of pesticide in predetermined

areas; however, these capabilities are primarily used to hold a targeted rate when equipment

speed changes. These same capabilities may be able to strategically target areas of production

fields with weeds if the exact location of individual weeds can be determined. A field map

containing the location of individual weeds may enable the site-specific actuation of flow valves

and spot-treat fields. To create a site-specific weed map, each weed must be located, and a

treatment boundary drawn. Site-specific maps would be necessary for each field multiple times

throughout the growing season for each weed generation. Only spraying targeted weeds in this

way is known as site-specific weed management (SSWM) and is almost non-existent because

site-specific maps are currently labor intensive to create, both in the field to locate weeds and in

defining treatment areas.

The convergence of three technologies may be able to automate site-specific weed map

generation. First, on-demand access to high-resolution field imagery from advanced sensors on

unmanned aerial systems (UAS) may provide the near-real time whole field information

necessary to capture and quantify weed presence. Second, the continued maturity and

commercial availability of machine learning models are increasingly accessible in user-friendly

interfaces and may be able to automate image stream analysis. Finally, the recent industry

adoption of pulse-width modulated (PWM) spray equipment, equipped with individual nozzle

control, may be able to provide the plant level resolution necessary to carry out site-specific

applications. Together, these three technologies have the potential to make SSWM accessible

and routine for today's producers. An accurate site-specific workflow has the potential to reduce

the input cost and the environmental impacts of many herbicide treatments. The cumulative

reduction in inputs and environmental impact of multiple applications in each field across

applicable production systems would be substantial. This work proposes and evaluates a novel

workflow leveraging this technological convergence to automate site-specific weed map

generation and application.

2.0 EVOLUTION OF SITE-SPECIFIC WORKFLOWS AND OPPORTUNITIES IN WEED MANAGMENT

Abstract

Precision agriculture research has made many production practices more strategic, cost effective, and efficient; however, pesticides are still broadcast across whole fields with no consideration for target variability. Advancements in remote sensing, object detection via deep learning, and sophisticated field equipment may meet the foundational requirements for widespread adoption of site-specific pesticide applications. A workflow that combines the capabilities of these technologies and facilitates access to accurate treatment maps could dramatically reduce the volumes of pesticides applied, reduce input cost for producers, and reduce environmental impact of production systems. Furthermore, in the case of selective herbicides or application made in genetically tolerant crops, the precision demands are much lower and may be feasible in current production systems. Evaluation of a workflow incorporating these converging technologies would be valuable to ensure the integrity of a resulting applications efficacy.

2.1 Precision agriculture

Precision agriculture is a broad term that includes sensor technology, information systems, and enhanced equipment to maximize input efficiency by accounting for variability in agriculture production systems (Gebbers & Adamchuk, 2010). Primary managed agricultural inputs include soil nutrients applied at variable rates based on site-specific need (Basso et al., 2016), high-speed precision planting of seeds (Xue et al., 2019), irrigation water management (Monaghan et al., 2013), and rate-controlled pesticide applications (Luck et al., 2011). Although

rate-controlled pesticide application equipment is considered part of precision agriculture, onboard field computers work to hold application rates constant and vary the flow of pesticides only as equipment speed changes (Luck et al., 2011). This results in broadcast rates of pesticides uniformly applied throughout fields despite fluctuating equipment speeds. Currently the technology is only being used to account for in-field speed fluctuation and to shut off individual nozzles or groups of nozzles at field boundaries or in already applied areas (Fernández-Quintanilla et al., 2018). Unlike other precision agriculture technologies, true site-specific pesticide treatments are almost non-existent in the United States (Nowak, 2021). Successful site-specific pesticide applications would require three components: comprehensive field scouting (sensors), accurate pest identification (information systems), and site-specific capable sprayers (enhanced equipment) (Gebbers & Adamchuk, 2010).

López-Granados (2011) suggested that two primary limitations of widespread adoption of site-specific workflows include a lack of skill sets amongst farmers necessary to take advantage of machine learning (ML) and remote sensing advances, and a lack of accessibility and compatibility of these systems with current field equipment. A review of the adoption of precision agriculture technologies in developed countries published in 2021 did not include SSWM technologies, only introducing the subject with a mention of variable rate crop protection products (Nowak, 2021). However, the same survey lists the adoption examples of georeferenced field scouting via UAS and automated section control for pest spraying, which are foundational technologies for SSWM workflows, if weeds can be reliably extracted from the imagery. Machine learning techniques to object detection may be able to provide the automated detection of weeds within UAS imagery necessary for practical site-specific workflows. While ML approaches to automated pest detection and decision making are often viewed as black box

technologies (Toda & Okura, 2019), user accessibility to theses advanced algorithms is improving. Sophisticated open-source deep learning algorithms have begun to be used in commercially available software with simplified graphical user interfaces (GUIs) reducing the skills necessary to access these advanced detection tools. The combination of user-friendly remote sensing (RS) via UAS, accessibility to powerful deep learning decision tools in GUI interfaces, and the current adoption of application equipment that can apply at sub-meter resolutions holds promise for the development of pragmatic site-specific workflows. Not only do these advancements meet the foundational criteria suggested by Gebbers & Adamchuk (2010), their accessibility and compatibility with today's field equipment may reduce the hurdles of adoption described by López-Granados (2011).

2.2   Field scouting for weeds with UAS

Unmanned aerial systems are a relatively new data collection tool available to agricultural production (Tenhunen et al., 2019). Coupled with high-resolution sensors, these platforms can provide a new perspective to agriculture fields that was previously inaccessible due to cost or regulatory restrictions (Khanal et al., 2017). Compared with satellite platforms, UAS's can produce crop imagery with superior spatial and temporal resolution flexibility (Murugan et al., 2017). Additionally, commercially available UAS can provide imagery at these resolutions at a relatively low entry cost which has expedited adoption (Lambert et al., 2019).

Site-specific management of crops, nutrients, and pests have long been goals of precision agriculture (PA)(Praveen & Sharma, 2020). The ability to strategically place inputs as intra-field needs change has the potential to significantly save on operating cost and reduce environmental impact. If RS is used for data collection, two requirements must be met regarding the target

objects or phenomenon. First, availability of high enough spatial resolution imagery for the object or phenomenon to be detected and second, its precise location (Gómez-Candón et al., 2014). Access to and robust analysis of real-time or near real-time data required to inform these site-specific management decisions has been a major limiting factor in the development of practical PA workflows (Castillejo-González et al., 2014; Mintert et al., 2016). In particular, the spatial resolution required to detect intra-field objects such as weeds with enough detail to differentiate them from crops has been a major hurdle in SSWM workflows leveraging remotely sensed data (Gómez-Candón et al., 2014).

Spatial resolution is the detail or sharpness that can be captured as a function of a sensor's instantaneous field of view (IFOV) and its observation height or altitude above ground level (AGL). Spatial resolution is often reported as its ground sample distance (GSD) which describes the nominal area of the ground that one pixel represents (i.e., 3 m, 3 cm).

Very high spatial resolution imagery (< 1 cm GSD) significantly aids in the high accuracy detection of relatively small objects such as emerging weeds in crop fields (Räsänen et al., 2019). Even with today's high megapixel sensors, low-altitude flights are often necessary to achieve the spatial resolution for robust detections (Gómez-Candón et al., 2014). However, low altitude imagery collection coupled with today's UAS batteries greatly limits the area that can be realistically imaged in a series of flights (Murugan et al., 2017). Therefore, determining a pragmatic balance between high accuracy detections at the lowest resolution requirements would be foundational to SSWM workflows using UAS derived imagery (Gómez-Candón et al., 2014).

2.3   Weed identification with deep learning models

Traditionally, pests are identified through manual and regular field scouting. For site-specific treatment at the plant-level resolution to be practical, targeted pests must be accurately identified and their location known (Lottes et al., 2020). In the case of whole field imagery captured by UAS, manual annotation of pests in imagery to produce site-specific spray maps is laborious and therefore not practical at today's farm scale.

Interactive supervised classification workflows have been demonstrated that were trained on individual fields imagery. Castaldi et al. (2017) used a support vector machine (SVM) assisted supervised classification approach to extract weed patches in corn fields using NDVI collected via UAS (35 m alt, 9 cm GSD). Actual field imagery was used for training, samples of soil, crop, and weed patches were hand annotated as training data and the ENVI 5.2 software package completed classification of the rest of the image with 61% accuracy.

Taking advantage of spectral differences between targets is one primary mechanism for image segmentation and classification (Fernández-Quintanilla et al., 2018; Louargant et al., 2017). Louargant et al. (2017) tested the classification accuracy of an unsupervised k-means algorithm (Mahalanobis distance) using UAS imagery (50 m alt, 6 cm GSD). Results suggest the detection of vegetation was relatively easy (95% accuracy) if pixels contained 60% or greater vegetation. However, spatial resolution was reported as being a limitation when discriminating between monocots and dicots (80% vegetation in pixel required for 80% accuracy). Methods that rely exclusively on spectral reflectance values have not provided the robust weed extraction necessary for real field applications to be generalized (Louargant et al., 2017). Many spectral classification workflows have additionally considered the location of vegetative pixels to increase accuracy in discriminating weeds from crops (Fernández-Quintanilla et al., 2018). Using

RGB imagery from UAS at ten meter altitude, Tenhunen et al. (2019) demonstrated a "row recognizer" workflow that used a k-means algorithm to extract vegetation pixels from soil using spectral values, followed by a Fourier transform that located the center of gravity of plant pixel clusters. Least squares regression was then applied to identify the number and orientation of planted rows in each frame, resulting in 94% accuracy at row detection. In a field application, any plant pixel groups outside of rows could be classified as weeds based on location alone.

Weeds that are within rows provides additional challenges. Plant leaf shape and structure have been explored in extracting crop occluded weeds. Local invariant features identify key points on the structure of target like weed species and are resistant to misclassifications caused by deviations in image rotation and altitude (Pflanz et al., 2018). Using a Bag of Visual Word (BoVW) approach (example of local invariant feature), Pflanz et al. (2018) were able to generate species specific individual plant weed maps at 89-94% accuracies using a UAS at one and six meter flight altitude. While impressive, the spatial coverage of today's UAS at these altitudes is excessively limiting, which lends the practical application of this model to proximal or ground-based methods only.

Workflows discussed thus far, except for Pflanz et al. (2018) primarily extract and classify on a pixel-by-pixel basis known as pixel-based image analysis (PBIA). De Castro et al. (2017) demonstrated an Object-based image analysis (OBIA) technique to map *Cynodon dactylon* in vineyards using UAS imagery. Red, green, and blue bands were segmented and resulting objects classified as soil, crop, and weeds with an overall 85% accuracy. A similar contribution by De Castro et al. (2018) applied an OBIA workflow to mapping weeds in sunflower and cotton fields using UAS imagery (30 m and 60 m alt, 0.6 and 1.2 cm GSD).

Again, RGB bands were segmented, and objects were classified with a random forest algorithm with 84% and 81% accuracy reported for weeds in cotton and sunflower crops, respectively.

More recently developed artificial neural networks (ANNs) have quickly shown utility in OBIA (Lin et al., 2019). Artificial neural networks have allowed for the fusion of textural information from imagery to spectral reflectance values which have further improved classification accuracy (Tamouridou et al., 2017). Tamouridou et al. (2017) used a multilayer perceptron ANN that received an image texture layer in addition to spectral reflectance bands to map pixels representing milk thistle (*Silybum marianum)* in UAS imagery (10 cm GSD) with 99.54% accuracy. Similarly, Pantazi et al. (2017) compared three counter-propagation ANN's (CP-ANN) in mapping *Silybum marianum* from UAS imagery (10 cm GSD) taken at 115 m altitude. Each model considered individual green, red, near-infrared bands, and a computed texture layer as inputs, resulting in 99%-pixel classification accuracy for each model.

Continued advances in computer vision in the form of convolutional neural networks (CNNs) has shown promising ability to quickly extract small features in imagery (Lin et al., 2019). Convolutional neural networks have proven themselves as one of the most successful architectures in modern deep learning methods at higher-level feature detecting and are replacing more conventional methods (Li et al., 2020). Furthermore, many open-source CNNs such as Faster R-CNN (Ren et al., 2017), Single Shot Detector (SSD), (Liu et al., 2016), RetinaNet (Lin et al., 2017), and You Only Look Once (YOLO) (Redmon et al., 2016; Redmon & Farhadi, 2018), have the ability to detect multiple classes of objects in the same image. The reported accuracies of these models, application in remotely sensed imagery, and ability to detect multiple classes at the same time make them appealing in geospatial workflows leveraging RS imagery (Pearse et al., 2020).

Modern object detection and classification CNN models are commonly categorized as single- and two-stage algorithms. Generally accepted as more accurate, two-stage models first create a Region Proposal Network (RPN) of potential targets and in a second step, classify the detections (Nguyen et al., 2020). These two-stage approaches provide for the best accuracy comparisons in benchmark competitions using published datasets like Pascal's Visual Object Classes (PASCAL-VOC) and Microsoft's Common Objects in Context (COCO); however, this accuracy often comes at the expense of training and detection speed (Ren et al., 2017).

Region proposed convolutional neural network (R-CNN) (Girshick et al., 2014), Fast R-CNN (Girshick, 2015), and Faster R-CNN (Ren et al., 2017) are a progressive series of high performing two stage, open-source object detection algorithms. Aptly named, each updated version has shown improvements in detection speeds with the goal of accurate real-time detections. The most recent version, Faster R-CNN first uses a series of feature extractors that segment images and aggregate segments at different scales to produce a proposed feature map (Figure 2.1). Next, it leverages a deep fully convolutional RPN that outputs areas of elevated interest within the proposed segmented feature map in the form of bounding boxes. Finally, the proposed boxes in the feature map are classified using the fully connected R-CNN detector (Ren et al., 2017).
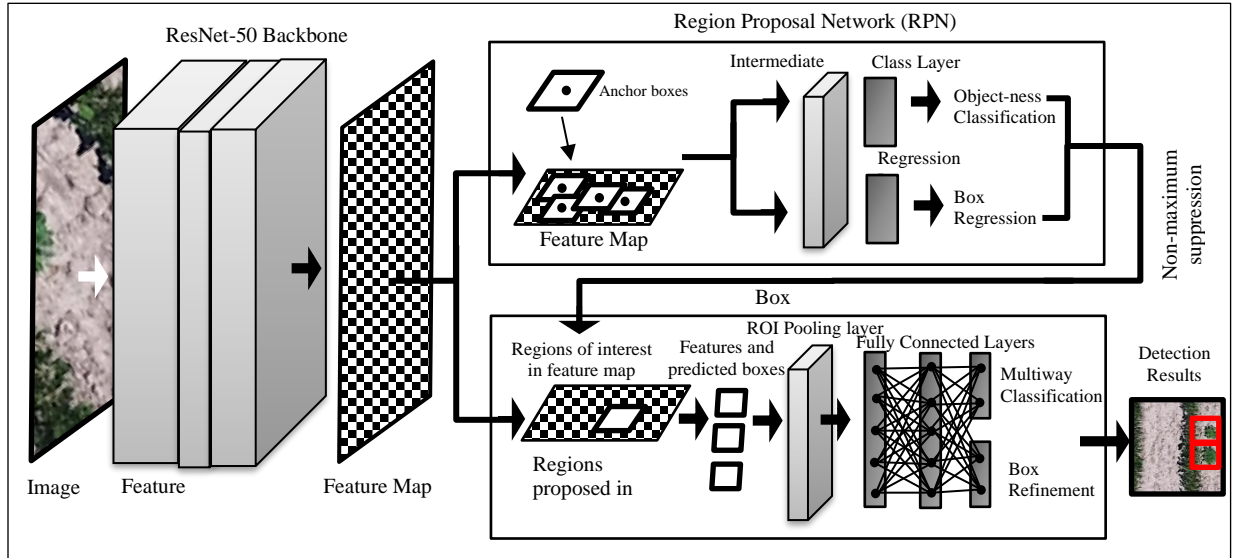
Figure 2.1: Convolutional layers structure of Faster R-CNN (Vasconez et al., 2020).

You Only Look Once (YOLO) was one of the first attempts at building a single stage object detection algorithm published by Redmon et al. in 2016 (Nguyen et al., 2020), followed by an updated version (YOLO v2) by the same authors in 2017. Redmon & Farhadi's published YOLO v3 in 2018 also an open source, single-stage object detection algorithm. YOLO v3 uses a feature extractor with 53 layers named by the authors "DarkNet-53" (Figure 2.2). Like its predecessors, predictions are made without using an RPN, or "only looking once" (Khan et al., 2021). It treats detections like a regression problem by dividing input images into grids and performing detections within each grid cell and then combining adjacent similar detections (Nguyen et al., 2020). YOLO v3 performs detections at three scale layers which enhances its ability to detect both large and small objects in imagery (Xu & Wu, 2020).
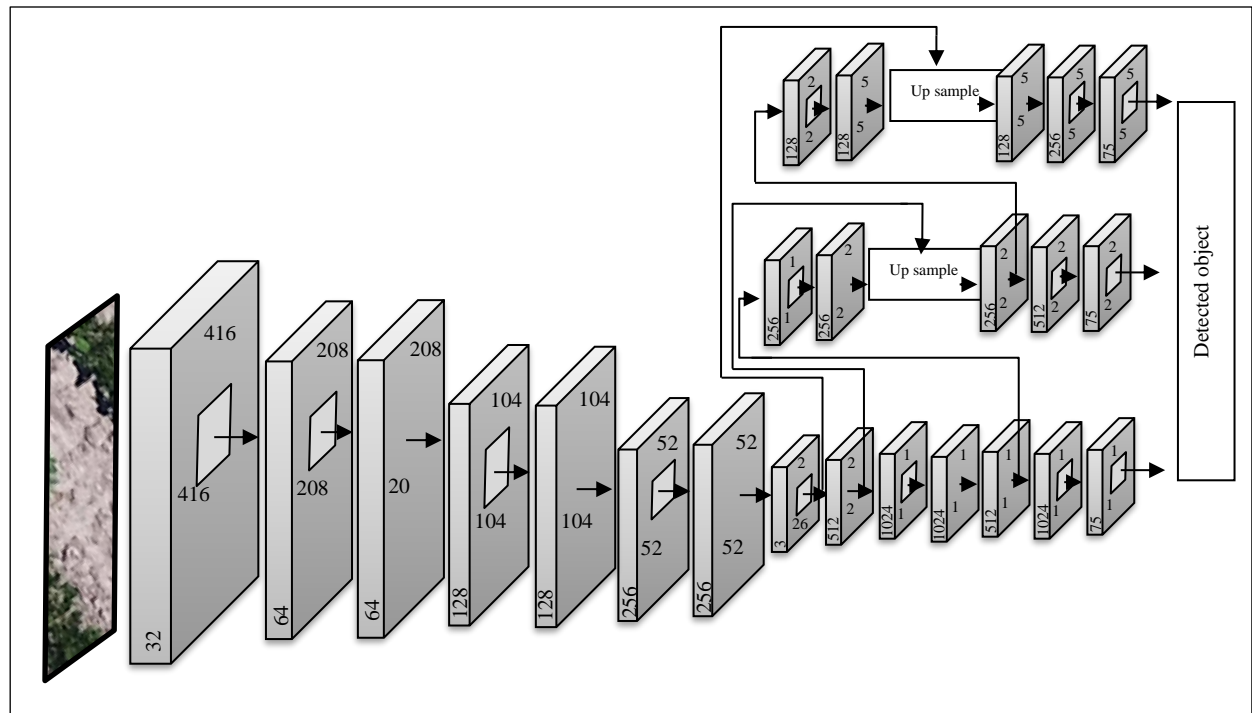
Figure 2.2: Convolutional layer structure of YOLO v3 (Tan et al., 2021).

Single Shot Detection (SSD) is another single stage detector which also integrates region proposal and classification into one step reducing the number of "looks" at imagery to one, therefore reducing detection speeds. Like Faster R-CNN, SSD first uses a feature extractor to segment and then aggregate segments at different scales (Lou et al., 2022). After feature extraction, SSD uses a pyramid structure of six convolutional layers to perform bounding box predictions and a regression analysis for feature classification in a single, continuous stage (Li et al., 2020) (Figure 2.3).
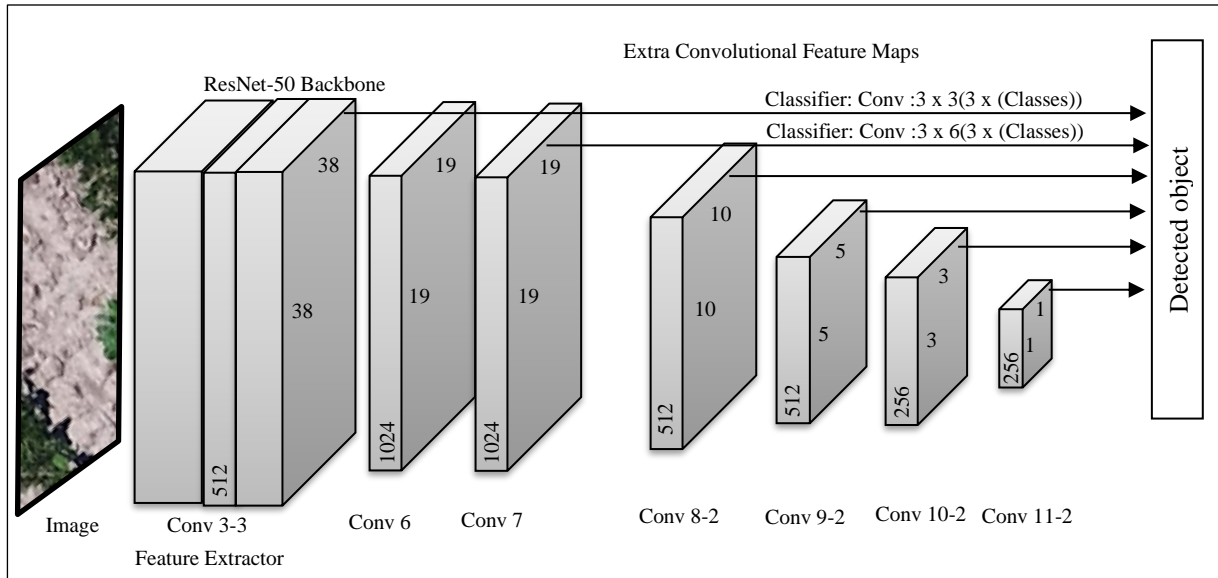
Figure 2.3: Convolutional layer structure of Single Shot Detector (Liu et al., 2016).

RetinaNet is another single stage object detection CNN published in 2017 by Lin et al. RetinaNet was designed to address one of the primary shortcomings of other single stage detection systems, the class imbalance between detections in the extreme foreground and background of images (Nguyen et al., 2020). This issue was addressed by the addition of focal loss to the classification "subnet" which weighted foreground detections different than the more difficult background detections during training (Lin et al., 2017). These weights focused learning on the more difficult detections and gained detection speeds by spending less time on foreground objects when deployed. RetinaNet uses a similar FPN as previously mentioned single stage detectors; however, horizontal connections from the feature extractor are connected to a convolutional layer in the models FPN (Figure 2.4).
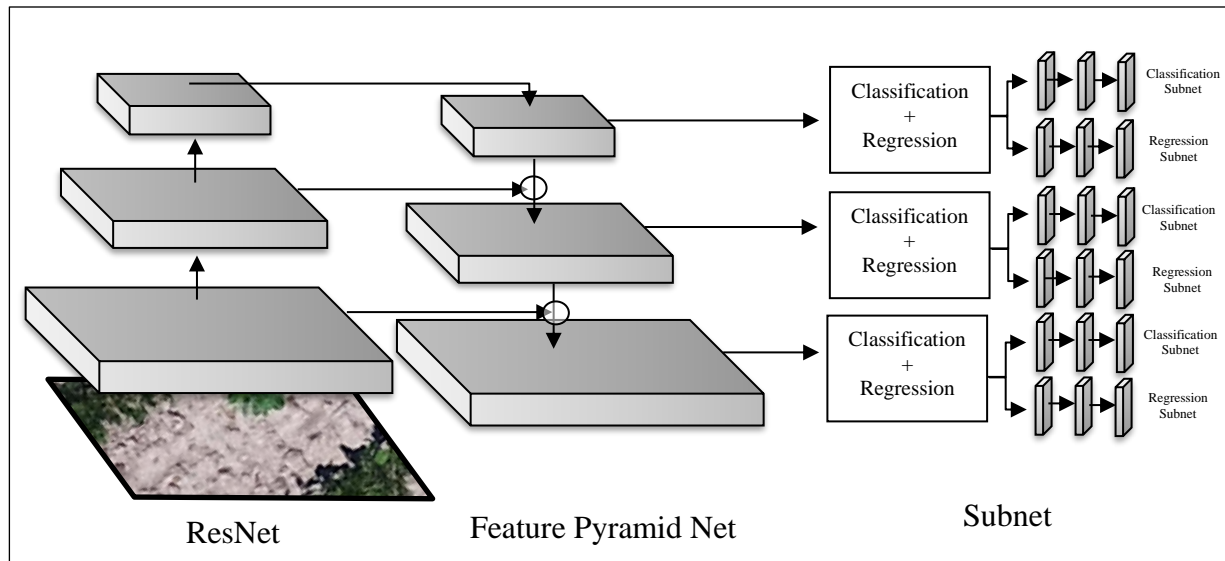
Figure 2.4: Convolutional layer structure of RetinaNet (Lin et al., 2017).

Recent open-source access to advanced object-detection CNNs such as Faster R-CNN, YOLO v3, SSD, and RetinaNet provide opportunities for their evaluation in plant level detections. Robust detections from models such as these could serve as the individual plant scouting step necessary to inform site-specific weed map generation.

2.4   Site-specific enabled equipment

Currently, site-specific spray equipment is available as manned equipment with actuators and real-time non-imaging sensors (Fernández-Quintanilla et al., 2018). Sensors are mounted across the spray boom and wired in series with spray nozzle actuators. Sensors are non-imaging scanners that are designed to sense and trigger action from equipment in real-time; however, these sensors simply differentiate between the reflectance curves of soil and plants (Fernández-Quintanilla et al., 2018). Although they can detect the presence of plants, they have no ability to distinguish between plant species, limiting their use to those pesticide applications that take place

prior to the emergence of the crop and excludes all applications post crop emergence (Souza et al., 2020). True site-specific treatment cannot be achieved currently with this form of equipment. This limitation has led to very poor adoption rates of this technology among production agriculture in the United States (Fernández-Quintanilla et al., 2018).

Accurate discrimination of weeds from crops is critical to in-crop site-specific treatments, and it is equally important that the application equipment be able to act on this information. Much research has focused on combining these two steps in advanced autonomous robotics known as "See and Spray" equipment that contain on-board decision capabilities and actuators. Gai et al. (2020) demonstrated the in-field real-time classification of weeds among broccoli and lettuce crops using a CNN algorithm fed color imagery and depth data. The CNN could identify crops and treat adjacent areas with an overall 91% accuracy at 3.2 km/hr. Lottes et al. (2020), tested their proposed CNN also in real-time on several versions of Bonirob multi-use autonomous agriculture robots in sugar beet fields. Further work was done more recently by Lottes et al. (2020) with the same field equipment and an updated algorithm. They demonstrated the ability to extract and classify different plant features including leaves and stems. The robotic and tractor mounted "See and Spray" systems mentioned have made impressive progress using deep learning algorithms to accurately classify crops/weeds and have improved software for a faster and more agile operation (Fernández-Quintanilla et al., 2018) However, while using real-time object detection via CNNs to inform advanced robotics on treatment areas is impressive, the scale and speed in which this equipment currently operates is limiting. Modern sprayers travel between 24 and 32 kph with 36-meter-wide spray booms resulting in a treatment rate of 14,400-19,200 m^2/min. The autonomous field robots as reported in Gai et al. (2020) travel closer to 3 kph and cover only what they can image and process. This would result in an estimated treatment

rate of 25 m^2/min assuming 0.5-meter-wide imaging window. This rate may be acceptable in some small specialty crop production systems but is not suitable for the thousands of hectares managed by most of today's row-crop producers. For this technology to be used in today's crop production systems, the processing speeds of onboard software must be increased, and the machinery must be scaled.

Selective treatment of pesticides may be possible with current commercially available equipment. Campos et al. (2020) were able to develop field zone specific spray maps that were used on a commercial sprayer. They varied the pesticide rate appropriately across the field as needs varied based on predetermined field zones. Similar control of the target pest was reported with reduced rates of water volume and total pesticides applied. Castaldi et al. (2017) used a commercially available sprayer informed with imagery collected from a UAS to develop spray maps and spot treat patches of weeds in corn fields. Their results showed nearly a 40% decrease in herbicide used when selectively spraying with no significant difference in the crop or weed biomass compared to the broadcast treatment. Both Campos et al. (2020) and Castaldi et al. (2017) were able to create "treatment zones" and "weed patch" versions of site-specific spray maps that adjusted the rate of pesticide sprayed as site-specific needs within fields changed. Furthermore, both were applied with equipment common to today's production systems. Castaldi et al. (2017) concluded that the efficiency of site-specific treatment was limited by the spatial resolution of the prescription map and the spatial resolution of the sprayer used. The equipment used in Castaldi et al. (2017) had an equipment resolution limited by two-meter boom sections leading them to classify treated areas as 'patch' treatments. Commercially available equipment now commonly has individually controlled nozzles on 50 cm spacing, which more closely approaches plant-by-plant site-specific capable applications.

The relatively recent adoption of pulse width modulated (PWM) spray technology in the agriculture industry has increased the accuracy and precision of pesticide rates and uniformity (Luck et al., 2011). PWM sprayers use highly responsive solenoids to turn on and shut off each nozzle at a duty cycle proportional to required flow rate (Butts, Luck, et al., 2019; Luck et al., 2011). This gives PWM equipment the ability to adjust rate as low as individual 50 cm nozzle spacing. Individual nozzle control may be able to vary rates of products as needed across fields or to selectively treat areas as needed. Leveraging the high spatial resolution capabilities of recently adopted PWM sprayers may provide the equipment requirements needed for higher resolution site-specific applications.

2.5   SSWM opportunities in today's production systems

Most of the post-emergent herbicides used commercially today are either selective or applied to herbicide tolerant (HT) crops (Meyer & Norsworthy, 2019). Selective herbicides include those that are only active on one type of plant with little or no adverse effects to other plants. Examples would be applying herbicide that controls only grasses to a broadleaf crop or applying a broadleaf herbicide to a monocot crop such as rice or corn. Genetically modified crops are also widely used in commercial production systems in the United States, including soybeans (Vialou et al., 2008). These crops are modified in a way that provides tolerance to certain herbicides, enabling the herbicide to be applied "over the top" of the crop canopy controlling only the target weeds (Bonny, 2016). First introduced in the 1990's with glyphosate tolerant soybeans (Green, 2009), additional crops, herbicide tolerances, and even stacked (multi-trait) tolerances are widely used in soybean fields in the United States and abroad (Bonny, 2016). While there is opportunity for savings on pesticide cost and reduced environmental impact with

17

SSWM workflows, the lack of consequences for applying these herbicides to the crop greatly reduces the accuracy demand and equipment resolution of selective and genetically tolerant herbicide applications.

A SSWM workflow should be evaluated in the context of broadcast application efficacy expectations and anticipated improvements in application precision due to reductions in treatment areas and pesticide inputs. Since broadcast applications treat the entire field, all weeds are targeted. While all weeds are targeted, limitations of equipment, the application environment, and pesticide chemistries often inhibit 100% control. Therefore, for this research, a practical threshold was considered for positive weed detections approaching broadcast expectations while acknowledging some room for error or missed detections. Additionally, broadcast application precision is negatively influenced by the areas void of weeds that are unnecessarily treated. Reducing any unnecessary treatment areas therefore improve application precision. Optimized application precision would minimize the treatment of areas void of weeds while not inadvertently missing detections. Thresholds for evaluation metrics were therefore chosen in the context of practical workflow adoption of equal weed control and conservative improvements in application precision over currently practiced broadcast applications.

2.6   Conclusion

Using very high-resolution imagery collected from UAS platforms may provide the complete field scouting needed for SSWM. Weed identification within this imagery has been demonstrated with today's publicly available deep learning models once properly trained. PWM application equipment can respond at a relatively fine resolution. The convergence of these three technologies may meet the foundational requirements for widespread adoption of site-specific

18

pesticide applications as described in Gebbers and Adamchuk (2010). SSWM may be possible with today's equipment but one major limitation is a lack of robust, reliable, and user-friendly workflows (Sishodia et al., 2020). A workflow that combines the capabilities of these technologies and facilitates access to accurate treatment maps would dramatically reduce the volumes of post-emergent herbicides applied, reduce input cost for producers, and reduce environmental impact of weed management in production systems. Furthermore, in the case of selective herbicides or application made in genetically tolerant crops, the precision demands are much lower and may be feasible in current production systems. However, a workflow incorporating these converging technologies must be evaluated to ensure the integrity of a resulting applications efficacy. These evaluations should include weed detection accuracy of different CNN models, consideration for the spatial resolution requirements to meet acceptable accuracies, and the application efficacy resulting from treatments applied by the high-resolution capabilities of today's application equipment.

2.7   Significance of this dissertation

This dissertation seeks to make two contributions to the field. A new high resolution dataset of *Palmer amaranth* in soybeans will be collected and annotated, one of the most common crops grown and challenging weeds to control in the Mississippi River Embayment. One of the primary limitations to applications of novel algorithms is adequate training data (Wang et al., 2019). This annotated dataset will be made available to be used for future advancements in object identification algorithms and potentially draw innovations from other industries into making application in this crop production system. Secondly, if successful, an automated workflow will be published that receives UAS field imagery as input, geolocates and

19

identifies *Palmer amaranth*, and exports a site-specific treatment layer that can be read by currently adopted field equipment. To the researcher's knowledge, a workflow such as this is not available and can act as a foundation to improve or expand upon with other algorithms, target pests, and other crops in future research.

Weed control remains one of the most costly and labor-intensive activities in modern crop production systems. The current practice of broadcast applications of pesticides to control weeds is expensive, often wasteful, and in some cases can negatively affect the environment. The successful creation of a site-specific treatment workflow could greatly reduce the input cost of producing food, feed, and fiber while reducing the environmental impacts of pesticide treatments. The cumulative effect of only applying pesticides to target organisms scaled across multiple applications, products, fields, and applicable production systems would be substantial.

2.8   Research objectives

The main goal of this dissertation research is to determine the prerequisites and limitations of a site-specific weed management workflow that leverages current UAS, deep learning models, and PWM spray technologies.

The first objective is to evaluate the *Palmer amaranth* detection accuracy of open-source CNNs (RetinaNet, Faster R-CNN, SSD, and YOLO v3) trained on imagery of coarsening spatial resolution, measured in ground sample distance (GSD). Detection accuracy was to be assessed in the context of practical workflow adoption of equal weed control and conservative improvements in application precision over currently practiced broadcast applications.

The second objective is to compare the control of *Palmer amaranth* in soybean plots between broadcast herbicide treatment and a SSWM treatment informed by a UAS imagery, weeds detected by a CNN, and applied by a PWM sprayer.

The hypotheses in this study are:

Ho$_1$ None of 16 custom trained open-source models (4 CNNs x training imagery of 4 spatial resolutions) will achieve $\geq$ 0.80 precision results at *Palmer amaranth* detection in soybean field imagery.

Ho$_2$ None of 16 custom trained open-source models (4 CNNs x training imagery of 4 spatial resolutions) will achieve $\geq$ 0.95 recall results at *Palmer amaranth* detection in soybean field imagery.

Ho$_3$ None of 16 custom trained open-source models (4 CNNs x training imagery of 4 spatial resolutions) will achieve $\geq$ 0.90 F1-score results at *Palmer amaranth* detection in soybean field imagery.

Ho$_4$ None of 16 custom trained open-source models (4 CNNs x training imagery of 4 spatial resolutions) will achieve a mean detection rate of 4 ha/h.

Ho$_5$ There are no significant ($p < .05$) differences in the mean treatment area (%) between plots sprayed with a broadcast rate and plots sprayed using prescription map generated by a SSWM workflow.

Ho$_6$ Plots selectively sprayed using prescription maps generated by a SSWM workflow will not achieve statically ($p < .05$) similar mean weed control compared with broadcast applications.

2.9   References

Basso, B., Fiorentino, C., Cammarano, D., & Schulthess, U. (2016). Variable rate nitrogen fertilizer response in wheat using remote sensing. *Precision Agriculture*, *17*(2), 168–182. https://doi.org/10.1007/s11119-015-9414-9

Bonny, S. (2016). Genetically Modified Herbicide-Tolerant Crops, Weeds, and Herbicides: Overview and Impact. *Environmental Management*, *57*(1), 31–48. https://doi.org/10.1007/s00267-015-0589-7

Butts, T. R., Luck, J. D., Fritz, B. K., Hoffmann, W. C., & Kruger, G. R. (2019). Evaluation of spray pattern uniformity using three unique analyses as impacted by nozzle, pressure, and pulse-width modulation duty cycle. *Pest Management Science*, *75*(7), 1875–1886. https://doi.org/10.1002/ps.5352

Campos, J., Gallart, M., Llop, J., Ortega, P., Salcedo, R., & Gil, E. (2020). On-Farm Evaluation of Prescription Map-Based Variable Rate Application of Pesticides in Vineyards. *Agronomy*, *10*(1), 102. https://doi.org/10.3390/agronomy10010102

Castaldi, F., Pelosi, F., Pascucci, S., & Casa, R. (2017). Assessing the potential of images from unmanned aerial vehicles (UAV) to support herbicide patch spraying in maize. *Precision Agriculture*, *18*(1), 76–94. https://doi.org/10.1007/s11119-016-9468-3

Castillejo-González, I. L., Peña-Barragán, J. M., Jurado-Expósito, M., Mesas-Carrascosa, F. J., & López-Granados, F. (2014). Evaluation of pixel- and object-based approaches for mapping wild oat (Avena sterilis) weed patches in wheat fields using QuickBird imagery for site-specific management. *European Journal of Agronomy*, *59*, 57–66. https://doi.org/10.1016/j.eja.2014.05.009

de Castro, A. I., Peña, J. M., Torres-Sánchez, J., Jiménez-Brenes, F., & López-Granados, F. (2017). Mapping Cynodon dactylon in vineyards using UAV images for site-specific weed control. *Advances in Animal Biosciences*, *8*(2), 267–271. https://doi.org/10.1017/S2040470017000826

de Castro, A., Torres-Sánchez, J., Peña, J., Jiménez-Brenes, F., Csillik, O., & López-Granados, F. (2018). An Automatic Random Forest-OBIA Algorithm for Early Weed Mapping between and within Crop Rows Using UAV Imagery. *Remote Sensing*, *10*(3), 285. https://doi.org/10.3390/rs10020285

Fernández-Quintanilla, C., Peña, J. M., Andújar, D., Dorado, J., Ribeiro, A., & López-Granados, F. (2018). Is the current state of the art of weed monitoring suitable for site-specific weed management in arable crops? *Weed Research*, *58*(4), 259–272. https://doi.org/10.1111/wre.12307

Fountas, S., Blackmore, S., Ess, D., Hawkins, S., Blumhoff, G., Lowenberg-Deboer, J., & Sorensen, C. G. (2005). Farmer Experience with Precision Agriculture in Denmark and the US Eastern Corn Belt. *Precision Agriculture*, *6*(2), 121–141. https://doi.org/10.1007/s11119-004-1030-z

Gai, J., Tang, L., & Steward, B. L. (2020). Automated crop plant detection based on the fusion of color and depth images for robotic weed control. *Journal of Field Robotics*, *37*(1), 35–52. https://doi.org/10.1002/rob.21897

Gebbers, R., & Adamchuk, V. I. (2010). Precision Agriculture and Food Security. *Science*, *327*(5967), 828–831. https://doi.org/10.1126/science.1183899

Girshick, R. (2015). Fast R-CNN. *2015 IEEE International Conference on Computer Vision (ICCV)*, 1440–1448. https://doi.org/10.1109/ICCV.2015.169

Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 580–587. https://doi.org/10.1109/CVPR.2014.81

Gómez-Candón, D., De Castro, A. I., & López-Granados, F. (2014). Assessing the accuracy of mosaics from unmanned aerial vehicle (UAV) imagery for precision agriculture purposes in wheat. *Precision Agriculture*, *15*(1), 44–56. https://doi.org/10.1007/s11119-013-9335-4

Green, J. M. (2009). Evolution of Glyphosate-Resistant Crop Technology. *Weed Science*, *57*(1), 108–117. https://doi.org/10.1614/WS-08-030.1

Khan, S., Tufail, M., Khan, M. T., Khan, Z. A., & Anwar, S. (2021). Deep learning-based identification system of weeds and crops in strawberry and pea fields for a precision agriculture sprayer. *Precision Agriculture*, *22*(6), 1711–1727. https://doi.org/10.1007/s11119-021-09808-9

Khanal, S., Fulton, J., & Shearer, S. (2017). An overview of current and potential applications of thermal remote sensing in precision agriculture. *Computers and Electronics in Agriculture*, *139*, 22–32. https://doi.org/10.1016/j.compag.2017.05.001

Lambert, J., Childs, D., & Freckleton, R. (2019). Testing the ability of unmanned aerial systems and machine learning to map weeds at subfield scales: A test with the weed *ALOPECURUS MYOSUROIDES* (Huds). *Pest Management Science*, ps.5444. https://doi.org/10.1002/ps.5444

Li, M., Zhang, Z., Lei, L., Wang, X., & Guo, X. (2020). Agricultural Greenhouses Detection in High-Resolution Satellite Images Based on Convolutional Neural Networks: Comparison of Faster R-CNN, YOLO v3 and SSD. *Sensors*, *20*(17), 4938. https://doi.org/10.3390/s20174938

Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollar, P. (2017). Focal Loss for Dense Object Detection. *2017 IEEE International Conference on Computer Vision (ICCV)*, 2999–3007. https://doi.org/10.1109/ICCV.2017.324

Lin, Wu, Fu, Wang, Zhang, & Kong. (2019). Dual-NMS: A Method for Autonomously Removing False Detection Boxes from Aerial Image Object Detection Results. *Sensors*, *19*(21), 4691. https://doi.org/10.3390/s19214691

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016). SSD: Single Shot MultiBox Detector. In *Computer Vision – ECCV 2016* (Vol. 1–Book, Section, pp. 21–37). Springer International Publishing. https://doi.org/10.1007/978-3-319-46448-0_2

López-Granados, F. (2011). Weed detection for site-specific weed management: Mapping and real-time approaches: Weed detection for site-specific weed management. *Weed Research*, *51*(1), 1–11. https://doi.org/10.1111/j.1365-3180.2010.00829.x

Lottes, P., Behley, J., Chebrolu, N., Milioto, A., & Stachniss, C. (2020). Robust joint stem detection and crop-weed classification using image sequences for plant-specific treatment in precision farming. *Journal of Field Robotics*, *37*(1), 20–34. https://doi.org/10.1002/rob.21901

Lou, X., Huang, Y., Fang, L., Huang, S., Gao, H., Yang, L., Weng, Y., & Hung, I.-K. uai. (2022). Measuring loblolly pine crowns with drone imagery through deep learning. *Journal of Forestry Research*, *33*(1), 227–238. https://doi.org/10.1007/s11676-021-01328-6

Louargant, M., Villette, S., Jones, G., Vigneau, N., Paoli, J. N., & Gée, C. (2017). Weed detection by UAV: Simulation of the impact of spectral mixing in multispectral images. *Precision Agriculture*, *18*(6), 932–951. https://doi.org/10.1007/s11119-017-9528-3

Luck, J., Sharda, A., Pitla, S., Fulton, J., & Shearer, S. (2011). A Case Study Concerning the Effects of Controller Response and Turning Movements on Application Rate Uniformity with a Self-Propelled Sprayer. *Transactions of the ASABE*, *54*(2), 423–431. https://doi.org/10.13031/2013.36445

Meyer, C. J., & Norsworthy, J. K. (2019). Influence of weed size on herbicide interactions for Enlist[TM] and Roundup Ready® Xtend® technologies. *Weed Technology*, *33*(4), 569–577. https://doi.org/10.1017/wet.2019.27

Mintert, J. R., Widmar, D., Langemeier, M., Boehlje, M., Erickson, B., Mintert, J. R., Widmar, D., Langemeier, M., Boehlje, M., & Erickson, B. (2016). *The Challenges of Precision Agriculture: Is Big Data the Answer?* https://doi.org/10.22004/AG.ECON.230057

Monaghan, J. M., Daccache, A., Vickers, L. H., Hess, T. M., Weatherhead, E. K., Grove, I. G., & Knox, J. W. (2013). More 'crop per drop': Constraints and opportunities for precision irrigation in European agriculture: More 'crop per drop.' *Journal of the Science of Food and Agriculture*, *93*(5), 977–980. https://doi.org/10.1002/jsfa.6051

Murugan, D., Garg, A., & Singh, D. (2017). Development of an Adaptive Approach for Precision Agriculture Monitoring with Drone and Satellite Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(12), 5322–5328. https://doi.org/10.1109/JSTARS.2017.2746185

Nguyen, K., Huynh, N. T., Nguyen, P. C., Nguyen, K.-D., Vo, N. D., & Nguyen, T. V. (2020). Detecting Objects from Space: An Evaluation of Deep-Learning Modern Approaches. *Electronics*, *9*(4), 583. https://doi.org/10.3390/electronics9040583

Nowak, B. (2021). Precision Agriculture: Where do We Stand? A Review of the Adoption of Precision Agriculture Technologies on Field Crops Farms in Developed Countries. *Agricultural Research*, *10*(4), 515–522. https://doi.org/10.1007/s40003-021-00539-x

Pantazi, X. E., Tamouridou, A. A., Alexandridis, T. K., Lagopodi, A. L., Kashefi, J., & Moshou, D. (2017). Evaluation of hierarchical self-organising maps for weed mapping using UAS multispectral imagery. *Computers and Electronics in Agriculture*, *139*, 224–230. https://doi.org/10.1016/j.compag.2017.05.026

Pearse, G. D., Tan, A. Y. S., Watt, M. S., Franz, M. O., & Dash, J. P. (2020). Detecting and mapping tree seedlings in UAV imagery using convolutional neural networks and field-verified data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *168*, 156–169. https://doi.org/10.1016/j.isprsjprs.2020.08.005

Pflanz, M., Nordmeyer, H., & Schirrmann, M. (2018). Weed Mapping with UAS Imagery and a Bag of Visual Words Based Image Classifier. *Remote Sensing*, *10*(10), 1530. https://doi.org/10.3390/rs10101530

Praveen, B., & Sharma, P. (2020). A review: The role of geospatial technology in precision agriculture. *Journal of Public Affairs*, *20*(1). https://doi.org/10.1002/pa.1968

Räsänen, A., Juutinen, S., Tuittila, E., Aurela, M., & Virtanen, T. (2019). Comparing ultra-high spatial resolution remote-sensing methods in mapping peatland vegetation. *Journal of Vegetation Science*, *30*(5), 1016–1026. https://doi.org/10.1111/jvs.12769

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788. https://doi.org/10.1109/CVPR.2016.91

Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. *ArXiv:1804.02767 [Cs]*. http://arxiv.org/abs/1804.02767

Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(6), 1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031

Sishodia, R. P., Ray, R. L., & Singh, S. K. (2020). Applications of Remote Sensing in Precision Agriculture: A Review. *Remote Sensing*, *12*(19), 3136. https://doi.org/10.3390/rs12193136

Souza, M. F. de, Amaral, L. R. do, Oliveira, S. R. de M., Coutinho, M. A. N., & Ferreira Netto, C. (2020). Spectral differentiation of sugarcane from weeds. *Biosystems Engineering*, *190*, 41–46. https://doi.org/10.1016/j.biosystemseng.2019.11.023

Tamouridou, A., Alexandridis, T., Pantazi, X., Lagopodi, A., Kashefi, J., Kasampalis, D., Kontouris, G., & Moshou, D. (2017). Application of Multilayer Perceptron with Automatic Relevance Determination on Weed Mapping Using UAV Multispectral Imagery. *Sensors*, *17*(10), 2307. https://doi.org/10.3390/s17102307

Tan, L., Huangfu, T., Wu, L., & Chen, W. (2021). Comparison of RetinaNet, SSD, and YOLO v3 for real-time pill identification. *BMC Medical Informatics and Decision Making*, *21*(1), 324. https://doi.org/10.1186/s12911-021-01691-8

Tenhunen, H., Pahikkala, T., Nevalainen, O., Teuhola, J., Mattila, H., & Tyystjärvi, E. (2019). Automatic detection of cereal rows by means of pattern recognition techniques. *Computers and Electronics in Agriculture*, *162*, 677–688. https://doi.org/10.1016/j.compag.2019.05.002

Toda, Y., & Okura, F. (2019). How Convolutional Neural Networks Diagnose Plant Disease. *Plant Phenomics*, *2019*, 1–14. https://doi.org/10.34133/2019/9237136

Vasconez, J. P., Delpiano, J., Vougioukas, S., & Auat Cheein, F. (2020). Comparison of convolutional neural networks in fruit detection and counting: A comprehensive evaluation. *Computers and Electronics in Agriculture*, *173*, 105348. https://doi.org/10.1016/j.compag.2020.105348

Vialou, A., Nehring, R. F., Fernandez-Cornejo, J., Grube, A., Vialou, A., Nehring, R. F., Fernandez-Cornejo, J., & Grube, A. (2008). *Impact of GMO Crop Adoption on Quality-Adjusted Pesticide Use in Corn and Soybeans: A Full Picture*. https://doi.org/10.22004/AG.ECON.6429

Wang, A., Zhang, W., & Wei, X. (2019). A review on weed detection using ground-based machine vision and image processing techniques. *Computers and Electronics in Agriculture*, *158*, 226–240. https://doi.org/10.1016/j.compag.2019.02.005

Xu, D., & Wu, Y. (2020). Improved YOLO-V3 with DenseNet for Multi-Scale Remote Sensing Target Detection. *Sensors*, *20*(15), 4276. https://doi.org/10.3390/s20154276

Xue, P., Xia, X., Gao, P., Ren, D., Hao, Y., Zheng, Z., Zhang, J., Zhu, R., Hu, B., & Huang, Y. (2019). Double-Setting Seed-Metering Device for Precision Planting of Soybean at High Speeds. *Transactions of the ASABE*, *62*(1), 187–196. https://doi.org/10.13031/trans.13055

3.0   DETECTION OF *PALMER AMARANTH* IN SOYBEAN IMAGERY USING OPEN-
SOURCE CNNS.

Abstract

Deep learning has the potential to become a prominent analysis tool in agricultural image processing applications for their abilities at finding objects such as weeds in imagery that are not practical for a human observer to detect at field scales. Automated and accurate detections could enable site-specific weed management (SSWM) workflows that selectively treat weeds amongst crops. High spatial resolution (0.4 cm) unmanned aircraft system (UAS) image streams were collected, annotated, and used to train 16 object detection convolutional neural networks (CNNs; RetinaNet, Faster R-CNN, Single Shot Detector, and YOLO v3 each trained on imagery with 0.4, 0.6, 0.8, and 1.2 cm spatial resolutions). Model detection precision, recall, and F1-scores were compared on imagery from four production fields containing approximately 7,800 weeds. Model performance improved for all CNN as spatial resolution of imagery increased. The highest performing model was Faster R-CNN trained on 0.4 cm imagery (precision = 0.86, recall = 0.98, and F1-score = 0.91).

3.1   Introduction

Autonomous robot technology has great potential to revolutionize agriculture production equipment by enabling automated decisions on the go. Spot-treatment of pesticides or manual spot-removal of pests is one example in production systems where "smart" equipment has obvious application. By exclusively treating the areas that are impacted by a particular pest, the technology would significantly reduce the overall use and waste of these pesticides. One of the

most critical components of an autonomous field system is its ability for precise detection and localization of pests or targets (Gai et al., 2020).

Today, weeds are identified through *in situ* field scouting, which has been established as laborious and often failing to capture the status of the entire field (Lambert et al., 2018; Varela et al., 2018). Once a weed 'flush' has been noted, a blanket treatment is prescribed to mitigate any competition between emerging weeds and the crop (Gianessi & Reigner, 2007; Sishodia et al., 2020). Overuse, environmental damage (Wang et al., 2019), and emerging resistance issues have called into question the broadcasting all herbicides and the goal of site-specific weed management (SSWM) has emerged as a desirable alternative for some applications. Site-specific weed management involves the targeted treatment of weeds within crops resulting in reduced total volume of herbicides applied and reduced environmental impact (Mahlein, 2016).

Research using proximal sensing of plant level weed targets with sensors mounted on ground-based equipment has found some commercial success at detecting green weeds amongst brown soil and crop residue (López-Granados, 2011). Known as green-on-brown detection, these examples exploit the spectral differences between living plants and non-plant material (Fernández-Quintanilla et al., 2018; Louargant et al., 2017). While real-time green-on-brown detection has been demonstrated to reduce chemical inputs, savings are limited to applications made prior to crop emergence, which has stifled their adoption (Fernández-Quintanilla et al., 2018). Additional work has focused on real-time proximal sensing of weeds amongst crops or green-on-green (Lottes et al., 2017, 2020; Wang et al., 2019). Robust green-on-green detection requires the ability to differentiate between plant species and therefore is a much more challenging problem. Recent advancements in computer vision, deep learning algorithms, and specifically convolutional neural networks (CNNs) have demonstrated impressive object

detection abilities and if trained properly, may be up to the challenge of species-specific plant detection in agricultural production fields.

Convolutional neural network's gain their namesake from the multiple hidden layers where "filters" are convolved over input images to extract different information such as textures, edges, and patterns (Boulent et al., 2019). One primary advantage of CNNs, is that the architectures allow for the spatial information within imagery to be retained making them well suited for spatial imagery problems like object detection (Lambert et al., 2019). Furthermore, they have been shown to do well at generalizing, with the ability to robustly detect features in data slightly different than their training sets (Boulent et al., 2019). Successful detection demonstrations have been made in plant disease detection (Toda & Okura, 2019), weed detection (Lambert et al., 2019; Milioto et al., 2018) and classifying nutrient deficiencies (Dobbels & Lorenz, 2019). Convolutional neural network's feature extraction precision is difficult to improve upon particularly when detecting small objects in remotely sensed imagery. Research has demonstrated the ability of CNNs to address the whole field identification need for site-specific treatments with image classification accuracies from 95-99% when properly trained (He et al., 2017; Liang et al., 2019; Lin et al., 2019). Recently developed and open-source access to advanced object-detection CNNs such as Faster R-CNN, YOLO v3, SSD, and RetinaNet has provided opportunities for their evaluation in custom trained use cases.

Nguyen et al. (2020) compared the performance of Faster R-CNN, SSD, YOLO v3, and RetinaNet at detecting vehicles in unmanned aircraft systems (UAS) and satellite imagery. Custom trained versions of each model were deployed on the VisDrone2019 dataset. Of these four CNNs, highest AP (IoU = 0.50) were achieved in car detections with RetinaNet (52.02), YOLO v3 (51.67), Faster R-CNN (46.85), and SSD (36.90).

Tan et al. (2021) evaluated the detection performance of RetinaNet, YOLO v3, and SSD at pill identification to assist pharmacist with unmarked medications. The custom trained RetinaNet model reported the highest mAP at 82.89%; however, detection rates were three times slower than the YOLO v3 model with a mAP of 80.69%. The real-time detection capabilities outweighed the slight accuracy differences and the authors reported that the YOLO v3 model would be a best fit for this use case.

Evaluations of custom trained versions of these models have also been performed in agricultural applications. Li et al. (2020) custom trained and compared a Faster R-CNN, YOLO v3 and an SSD model to detect agriculture greenhouses in rural China. Greenhouses (18,385) were annotated in two-meter satellite imagery and used as training data in each of the model architectures. YOLO v3 provided the best results reported a mAP of 90.4%, followed by Faster R-CNN and SDD with mAP's of 86.0% and 84.9%, respectively.

Lou et al. (2022) used Faster R-CNN, YOLO v3 and SSD to detect loblolly pine tree crowns in UAS imagery. Two sites, differing in tree maturity were used to evaluate the models with SSD performing the best at both sites (R^2 of 0.92 and 0.94 at the young and mature sites, respectively). YOLO v3 produced an R^2 of 0.62 and 0.69 at the young and mature sites, respectively. Faster R-CNN produced an R^2 value of 0.88 at the young site but failed to produce results at the mature site.

Vasconez et al. (2020) used a custom trained SSD and Faster R-CNN model to detect and count Avocado, apple, and lemon fruit on trees prior to harvest where 3,462 images of fruit

Khan et al. (2021) used 2,800 images (0.3 cm GSD) of weeds to custom train and compare Faster R-CNN and YOLO v3 at weed detection in strawberry and pea fields. The

average weed identification accuracy was reported at 91.0% for YOLO v3 and 85.6% for Faster R-CNN.

Many of the mentioned CNNs are designed as true real-time detections with frames per second metrics listed alongside reported accuracies. While these metrics are impressive and practical for many use cases, agriculture pesticide application equipment can cover 180 m^2/sec (27 m sprayer boom traveling at 6.7 m/s velocity). The process of capturing a single frame, deciding on its contents, relaying that decision to mechanical solenoids to then turn on nozzles in real-time at field speeds is a processing challenge for today's equipment. This is particularly true when scaled across common boom lengths (27 to 36 m). Improvements in onboard computer processing times are likely necessary before CNNs can be used for selective herbicide applications on-site and in real-time (Fernández-Quintanilla et al., 2018).

One potential solution is to use image streams from UAS equipped with high resolution RGB sensors to image a field prior to the application as part of *in situ* scouting. Automated weed detection could then be performed using a custom trained CNN on a separate, optimized computer platform alleviating the real-time processing bottleneck and the requirement of extensive onboard computing power.

While CNN's have demonstrated their ability in capturing weed conditions in UAS imagery to varying degrees better than observations made by traditional scouting techniques alone (Pantazi et al., 2017; Tamouridou et al., 2017); plant-level studies applying these modern CNNs to UAS imagery for site-specific decisions are lacking, and further research into weed mapping via UAS is warranted (Fernández-Quintanilla et al., 2018). For these workflows to be accurate and more practical for field applications, custom models are likely to be optimized at

detecting specific weeds in specific crops. Additionally, optimal spatial resolutions that balance detection accuracy and imagery collection efficiency must be determined.

This research will compare four open-source object detection CNNs (RetinaNet, Faster R-CNN, SSD, and YOLO v3) trained on imagery of increasing spatial resolution (0.4, 0.6, 0.8, and 1.2 cm GSD) to detect *Palmer amaranth* in *Glycine max* (soybean*)* production systems. *Palmer amaranth* is listed as one of the most troublesome weeds in production agriculture in North America according to multiple surveys conducted by the Weed Science Society of America (Van Wychen, 2016, 2017, 2021).

3.2   Training imagery collection, preprocessing, and annotation

3.2.1   Training imagery collection site

A series of three-meter by nine-meter plots were established at the Jackson County Extension Center (35.5715763°N, 91.2539575°W) located near Newport, Arkansas, USA (Figure 3.1). Each of the 104 soybean plots were either untreated or one of four replicated treatments that included 33 pre-emergent herbicide combinations. Plots were planted on June 15th, and pretreatments applied June 16th of 2021. These treatments produce plots with staggered *Palmer amaranth* emergence and varying weed density (plants/m^2). This induced variability was intended to mimic differences in *Palmer amaranth* size and populations found in production soybean fields.

Figure 3.1: Soybean plots with staggered *Palmer amaranth* emergence established to generate training samples.

3.2.2 Training imagery collection

Imagery was collected using a DJI Zenmuse X5s RGB sensor (iFlight Tech Company

Ltd. Shenzhen, China) with the DJI micro 4/3 (MFT) lens (iFlight Tech Company Ltd.

Shenzhen, China). This sensor and lens configuration has a 4/3" CMOS, 72° FOV, 15 mm focal

length, f/1.7 focal ratio, and focus to infinite. Image resolution of the camera is 5280×3956

(20.8 MP). The X5s sensor was mounted onto a DJI Inspire II (iFlight Tech Company Ltd. Shenzhen, China) quadcopter (Figure 3.2).

All UAS flights were conducted in class G airspace and by a licensed part 107 pilot in compliance with Federal Aviation Administration (FAA) regulations. Each unique collection event's flight plan was built within the mobile application Pix4D Capture 4.13.1-1 (Pix4D Inc. Lausanne, Switzerland) to facilitate image capture for processing in Pix4Dmapper 4.6.4 (Pix4D Inc. Lausanne, Switzerland). Overlap settings were 80% front and 70% side for each flight. During collection events, the UAS was manually controlled at take-off, ascended to target altitude (18.30 m), and navigated into the target area using the mobile application DJI GO 4 v4.3.46 (DJI Inc.). While hovering at altitude, the X5s's focus, and exposure was optimized for each flight using the automated procedure within the DJI application where the operator selects objects displayed on the controller screen and the software adjusts sensor aperture and focus accordingly. The DJI application was then closed, and the automated flight plan was initiated in the Pix4D Capture application. Training imagery was collected two and four weeks after soybean emergence when soybeans were approximately at V3 (three trifoliate leaves) and V7 (seven trifoliate leaves) growth stages, respectively.

Figure 3.2: DJI Inspire II quadcopter with Zenmuse X5s sensor and DJI MFT 15mm lens.

Table 3.1 shows the flight conditions while collecting the imagery used for training. Each flight was completed within two hours of local solar noon on their respective day to limit weed occlusion from shadowing. Altitude, flight velocity, image overlap, and GSD were held constant at 18.3 m, 1.83 m/s, and 0.4 cm. Average wind velocity was reported in 15-min increments from a Davis Vantage Pro2™ Premium Weather Station (Davis Instruments Corporation, Hayward, CA 94545 USA) that was located 270 m from collection sites. Wind velocity throughout each flight was stable at 1.56 and 2.24 m/s for flight one and two, respectively.

Table 3.1: Flight information for training imagery collection events.

Date	Time (CST)	Solar noon (CST)	Altitude (m)	Overlap (Front, Side)	Flight VEL (m/s)	GSD (cm)	Wind VEL (m/s)	Growth Stage
2021, July 8	11:17 to 11:26 h	13:10 h	18.30	80%, 70%	1.83	0.4	1.56	V3
2021, July 23	14:30 to 14:38 h	13:11 h	18.30	80%, 70%	1.83	0.4	2.24	V7

3.2.3 Training imagery preprocessing and annotation

Pix4Dmapper was used to co-register and orthorectify raw UAS images using the Ag RGB template. The rectified field orthomosaic (0.4 cm spatial resolution) was imported into ArcGIS Pro 2.8.2 (Esri inc. Redlands, California) for spatial analysis. Areas in the orthomosaic that corresponded to 100% pigweed cover (Figure 3.3), were removed from the orthomosaic because they were not representative of a field where a site-specific application would be applicable. The geoprocessing tool "Resample (Data Management Tools)" was applied to the 0.4 cm spatial resolution layer to produce three additional, identical, but reduced spatial resolution orthomosaics (0.6, 0.8, and 1.2 cm GSD).

Figure 3.3: Example area with 100% Palmer amaranth cover removed from training samples.

Artifacts (*Palmer amaranths* and broadleaf weeds) were labeled within the original (0.4 cm GSD) orthomosaic using the "Label Objects for Deep Learning Tool". Bounding boxes were placed around target weeds and given a class label either "pigweed" or "broadleaf." Object that could be definitively identified as *Palmer amaranth* were labeled as "pigweed" and definitive weeds that were too small, partially occluded, or blurred were given the generic label "broadleaf". This distinction is not necessary for field application since any positively identified weed should be treated; however, by creating and training with a very clearly defined "pigweed" class, a model likely would achieve higher accuracy when detecting *Palmer amaranth* within rows and amongst soybean plants. The "Label Objects for Deep Learning Tool" in ArcGIS Pro creates a vector layer of labeled and georeferenced bounding boxes can be associated with a raster layer. This layer, labeled on the highest resolution images set, was used as the input layer for each lower resolution raster in subsequent steps. This ensured that the labeled samples provided to each model were identical (i.e., number of labeled samples, framing of objects in bounding boxes, equally balanced classes) and only differing in the spatial resolution of the input raster. The "Export Training Data for Deep Learning Tool" was used to prepare imagery for training by dividing the rectified orthomosaic layer into uniform chips of pixels representing 2.5 x 2.5 meters of field imagery. Image chip size was selected to ensured that each labeled chip contained three rows of soybeans, providing the model with context for weed identification. Exported chip resolution by image set included 0.4, 0.6, 0.8, and 1.2 cm resolutions, representing 18.3, 26.3, 36.6, and 52.6 m flight altitude producing chips of 640 x 640, 480 x 480, 320 x 320, and 256 x 256 pixels, respectively. Image samples from each set are shown in Figure 3.4.
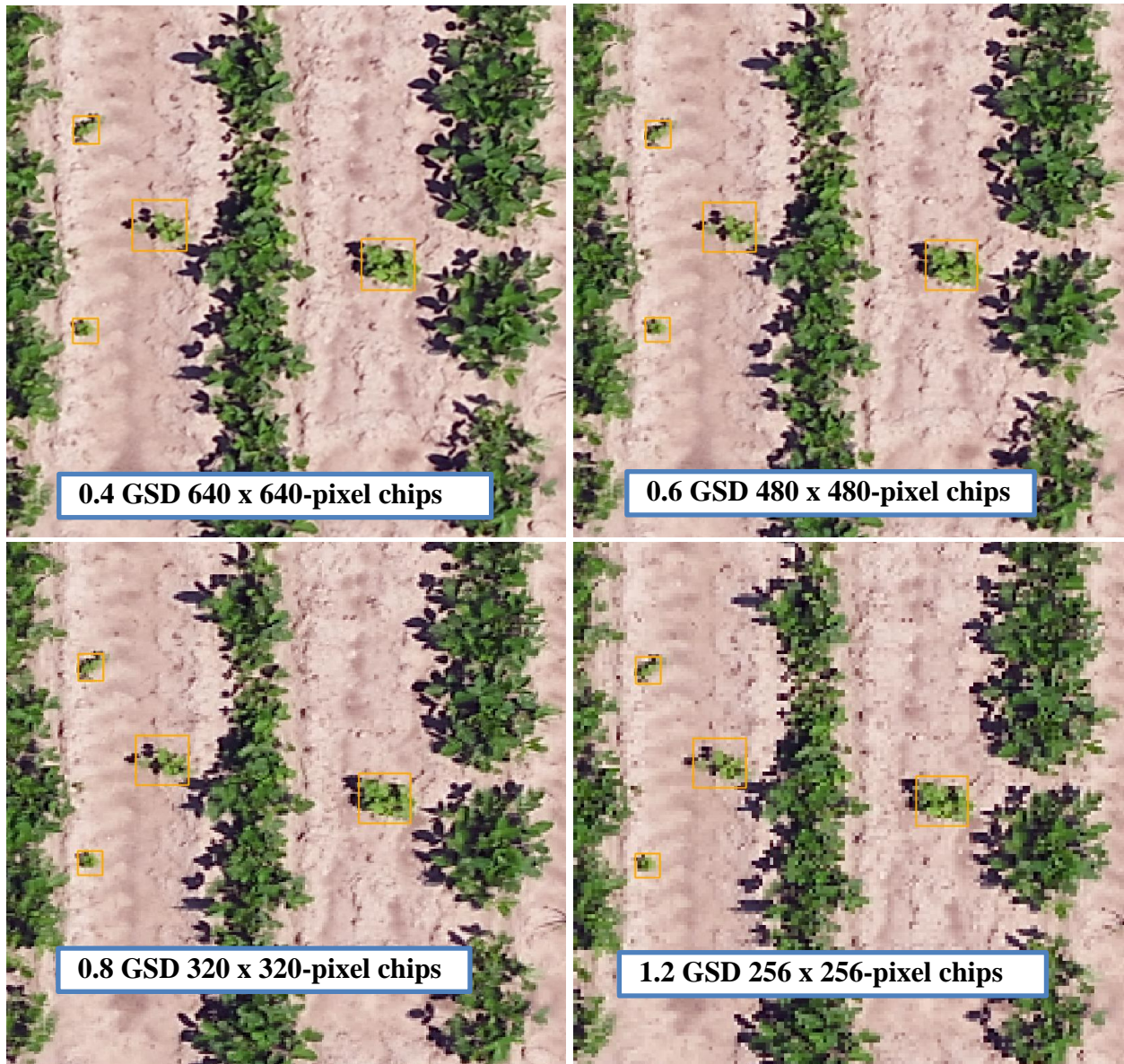
Figure 3.4: Image sets include 0.4, 0.6, 0.8, and 1.2 cm GSD representing 18.3, 26.3, 36.6, and 52.6 m flight altitude producing chips of 640 x 640, 480 x 480, 320 x 320, and 256 x 256 pixels respectively.

Training image chips were exported as TIFF files and associated metadata was exported in the PASCAL Visual Object Class format. This process was repeated for each of the four labeled orthomosaics, producing four training sets each containing 5,138 labeled weed samples represented in 0.4, 0.6, 0.8, and 1.2 cm spatial resolution imagery.

Table 3.2 summarizes mean, and standard deviation of the bounding boxes framed over each weed for each date and class. Additionally, the overall and class counts are reported. Figure 3.5 illustrates example training imagery at different spatial resolutions and bounding box labels, representing pigweed and broadleaf classes.

Table 3.2: Bounding box size and count information by training imagery flight. Bounding boxes tightly framed labeled weeds in training imagery and therefore correlated with weed size as represented by surface area in the orthomosaic.

	All Classes		Pigweed		Broadleaf	
	Count	Bounding box size (cm^2) Mean (SD)	Count	Bounding box size (cm^2) Mean (SD)	Count	Bounding box size (m^2) Mean (SD)
Flight 1	2509	300 (+/- 300)	1737	400 (+/- 400)	772	100 (+/- 100)
Flight 2	2629	734 (+/- 514)	2559	743 (+/- 515)	70	379 (+/- 300)

Figure 3.5: Example training sample data of decreasing spatial resolution. Note. Example training sample data from July 8th, 2021, of decreasing spatial resolution 0.4, 0.6, 0.8, and 1.2 cm GSD representing 18.3, 26.3, 36.6, and 52.6 m flight altitude. Pigweed and broadleaf are labeled with green and red bounding boxes, respectively.

3.3   Experimental Methods

This experiment compared the dependent variables' precision, recall, F1-score, and detection rate (ha/h) produced by independent variable "custom trained CNN" (CNNs: Faster R-CNN, SSD, RetinaNet, and YOLO v3 each trained on imagery of spatial resolutions 0.4, 0.6, 0.8, and 1.2 cm).

3.3.1 Manually annotated key assessment

The results of model accuracy testing, and subsequent field accuracy testing, were derived from comparisons made to manually annotated keys that were labeled using the highest spatial resolution (0.4 cm GSD) orthomosaic for each field/image set. An evaluation was conducted to assess the accuracy of detecting small weeds in these orthomosaic derived weed labels. A production field was imaged that contained a variation of weed sizes including a recent flush of very small weeds (< 2.5 cm height). RTK corrected GNSS ground reference location, weed size, and weed classification (pigweed and broadleaf) were collected for each weed through *in situ* field scouting on the date that imagery was collected. The weeds in the orthomosaic were labeled as described in section 3.3.2. This labeled layer was then compared to the *in situ* collected ground reference objects and true positive, false positive, and false negative results are reported. From these detection results, recall by weed size was assessed to evaluate the ability of UAS derived orthomosaics to model small weeds (< 2.5 cm height).

3.3.2 Model training

ArcGIS Pro was used to train each of the models through the "Train Deep Learning Model (Image Analysis tools)". Parameters for model training can be found in Table 3.3. The experimental platform configuration used for testing was the following: OS, Windows 10; GPU NVIDIA Quadro RTX 3000 with Max-Q Design; CPU, Intel® Core™ i7-1065 G7 CPU @ 1.30 GHz with 32 Gb of RAM. All 16 models were trained on this configuration. The orthomosaics used for training data were split into ninety percent by area for training and ten percent by area withheld for testing. Within the training imagery the data were split again 90 percent of the

43

training chips were used for training and ten percent were used for validation between epochs. Upon training completion, the results of the final validation set were reported for each model.

Table 3.3 summarizes the training parameters for the 16-custom trained CNNs evaluated. The same backbone model (Resnet-50) was used for CNNs RetinaNet, Faster R-CNN, and SSD. A different backbone model was used for YOLO v3 because it is exclusively compatible with Darknet-53. Both backbone models were trained on the ImageNet data set containing over one million images and were of similar size measured by their convolutional layers (50 and 53 layers for Resnet-50 and DarkNet-53, respectively). Image size (pixels) was adjusted proportionally with training imagery GSD so that each training image represented identical data or area in the field. Learning rate was automatically determined for each model within the training tool and is reported in Table 3.3. Batch size was limited by dedicated GPU memory on the training hardware and was held constant at two. Max epochs were set to 20.

Table 3.3: Training parameters for 16 custom trained CNNs.

Model Number	CNN	Backbone model	Image resolution (cm GSD)	Image size (pixels)	Learning Rate	Batch size	Max Epochs
1	RetinaNet	Resnet-50	0.4	640 x 640	0.00001, 0.0001	2	20
2	RetinaNet	Resnet-50	0.6	480 x 480	0.00004, 0.0004	2	20
3	RetinaNet	Resnet-50	0.8	320 x 480	0.00002, 0.0002	2	20
4	RetinaNet	Resnet-50	1.2	256 x 256	0.00002, 0.0003	2	20
5	Faster R-CNN	Resnet-50	0.4	640 x 640	0.000004, 0.00004	2	20
6	Faster R-CNN	Resnet-50	0.6	480 x 480	0.000004, 0.00004	2	20
7	Faster R-CNN	Resnet-50	0.8	320 x 480	0.000005, 0.00005	2	20
8	Faster R-CNN	Resnet-50	1.2	256 x 256	0.000004, 0.00004	2	20
9	SSD	Resnet-50	0.4	640 x 640	0.0002, 0.002	2	20
10	SSD	Resnet-50	0.6	480 x 480	0.0002, 0.002	2	20
11	SSD	Resnet-50	0.8	320 x 480	0,0004, 0.004	2	20
12	SSD	Resnet-50	1.2	256 x 256	0.0006, 0.006	2	20
13	YOLO v3	DarkNet-53	0.4	640 x 640	0.0002, 0.002	2	20
14	YOLO v3	DarkNet-53	0.6	480 x 480	0.00008, 0.0008	2	20
15	YOLO v3	DarkNet-53	0.8	320 x 480	0.0001, 0.001	2	20
16	YOLO v3	DarkNet-53	1.2	256 x 256	0.0003, 0.003	2	20

3.3.3   Model testing

Once trained each model was used to detect weeds on the ten percent of imagery withheld for testing. Testing parameters are summarized in Table 3.4. ArcGIS Pro's "Detect Objects Using Deep Learning (Image Analysis)" tool was used with detection thresholds set to 0.05. Padding was applied at ten percent of image pixel dimensions. Batch size was held constant at two. To gain a better measure of model accuracy, a Non-Maximum Suppression (NMS) algorithm was used as a post processing step to remove excessively overlapping, duplicate detections with an intersection over union (IoU) at or above 0.1. The "Compute Accuracy for Object Detection (Image Analyst Tool) was used to compare predictions made by each model to

the manually annotated key using an IoU of 0.5. True positive, false positive, and false negative detection results were organized and used to calculate precision, recall, and F1-scores.

Table 3.4: Model testing parameters.

Model Number	CNN	Image resolution (cm GSD)	Image size (pixels)	Batch size	Padding (pixels)	Detection threshold	NMS threshold (IoU)
1	RetinaNet	0.4	640 x 640	2	64	0.05	0.1
2	RetinaNet	0.6	480 x 480	2	48	0.05	0.1
3	RetinaNet	0.8	320 x 480	2	32	0.05	0.1
4	RetinaNet	1.2	256 x 256	2	26	0.05	0.1
5	Faster R-CNN	0.4	640 x 640	2	64	0.05	0.1
6	Faster R-CNN	0.6	480 x 480	2	48	0.05	0.1
7	Faster R-CNN	0.8	320 x 480	2	32	0.05	0.1
8	Faster R-CNN	1.2	256 x 256	2	26	0.05	0.1
9	SSD	0.4	640 x 640	2	64	0.05	0.1
10	SSD	0.6	480 x 480	2	48	0.05	0.1
11	SSD	0.8	320 x 480	2	32	0.05	0.1
12	SSD	1.2	256 x 256	2	26	0.05	0.1
13	YOLO v3	0.4	640 x 640	2	64	0.05	0.1
14	YOLO v3	0.6	480 x 480	2	48	0.05	0.1
15	YOLO v3	0.8	320 x 480	2	32	0.05	0.1
16	YOLO v3	1.2	256 x 256	2	26	0.05	0.1

3.3.4   Model Metrics

Trial size and model specific detection time (min) were recorded and used to calculate model-trial specific detection rates (ha/h). Precision and recall were calculated for each of the sixteen model results by comparing detections to the keyed layer detections using an IoU of 0.5 (Equation 3.1 and Figure 3.6).

Equation 3.1: Formula used to calculate Intersection over Union between ground reference layer and predictions.

$$\frac{Area\ of\ intersect}{Area\ of\ union} = \text{IoU}$$
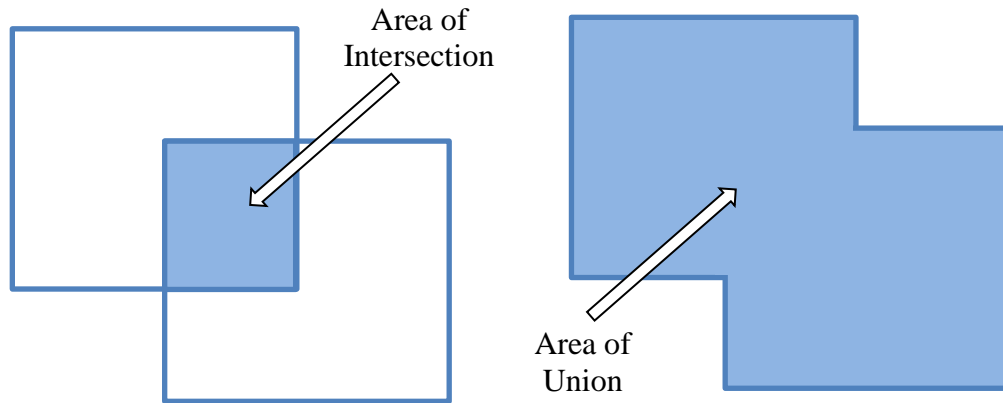
Figure 3.6: Visualization of IoU.

Precision was defined as the ratio of true positive detections of the positive detections made. While precision was important to reduce unnecessary treatments, it was considered indirectly as an input in calculating F1-scores for each model output. Precision was calculated using the Equation 3.2.

Equation 3.2: Precision equation

$$Precision = \frac{TruePositives}{TruePositives + FalsePositives}$$

Recall was defined by the ratio of true positive detections to true positives in the dataset including those not detected. Since the cost of not detecting weeds is much higher than

unnecessarily applying herbicides to crop, priority was given to models with high recall. Recall was calculated using Equation 3.3.

Equation 3.3: Recall equation

$$Recall = \frac{TruePositives}{TruePositives + FalseNegatives}$$

The harmonized mean of precision and recall is known as F1-score. Secondary to model recall, F1-score was used to evaluate model results using Equation 3.4.

Equation 3.4: F1-score calculation

$$F1 - score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

3.3.5   Field trial collection sites

Subsections of four soybean fields were used as replications for model deployment evaluation at the Jackson County Extension Center (35.5715763°N, 91.2539575°W) located near Newport, Arkansas, USA (Figure 3.7) in 2021 and 2022. Soybean rows in each field were planted at 76 cm spacing, soybean maturity differed slightly amongst replications, but all were middle to late vegetative growth stages (four to seven trifoliate leaves, V4-V7) entering reproductive stages (beginning flower, R1). *Palmer Amaranth* was the predominant weed species with size (height) spanning between emergent (in one field) and 60 cm.
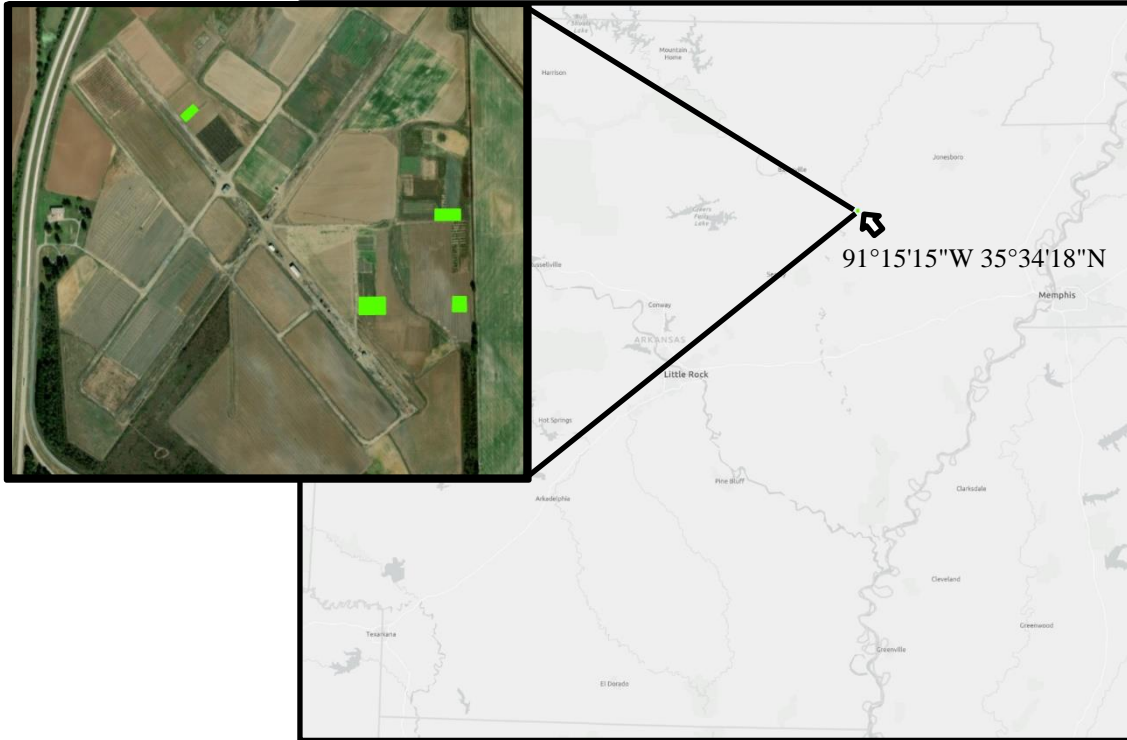
Figure 3.7: Locations of four field trials conducted in 2021 and 2022.

UAS imagery in each field was collected within two hours of local solar noon on their respective day to limit weed occlusion from shadowing. Altitude, flight velocity, image overlap, and GSD were held constant at 18.3 m, 1.83 m/s, and 0.4 cm. Imagery was collected using a DJI Zenmuse X5s RGB sensor (iFlight Tech Company Ltd. Shenzhen, China) with the DJI micro 4/3 (MFT) lens (iFlight Tech Company Ltd. Shenzhen, China). This sensor and lens configuration has a 4/3" CMOS, 72° FOV, 15 mm focal length, f/1.7 focal ratio, and focus to infinite. Image resolution of the camera is 5280 $\times$ 3956 (20.8 MP). The X5s sensor was mounted onto a DJI Inspire II (iFlight Tech Company Ltd. Shenzhen, China) quadcopter.

All UAS flights were conducted in class G airspace and by a licensed part 107 pilot in compliance with Federal Aviation Administration (FAA) regulations. Each unique collection

49

event's flight plan was built within the mobile application Pix4D Capture 4.13.1-1 (Pix4D Inc. Lausanne, Switzerland) to facilitate image capture for processing in Pix4Dmapper 4.6.4 (Pix4D Inc. Lausanne, Switzerland) where80% front overlap and 70% side overlap were used for each flight.

During collection events, the UAS was manually controlled at take-off, ascended to target altitude, and navigated into the target area using the mobile application DJI GO 4. While hovering at altitude, the X5s's focus, and exposure was optimized for each flight using the automated procedure within the DJI application where the operator taps on objects displayed on the controller screen and the software adjusts sensor aperture and focus accordingly. The DJI application was then closed, and the automated flight plan was initiated in the Pix4D Capture application.

3.3.6   Data analysis

True positive, false negative, and false positive detection results were aggregated from each of the four field deployments and used to calculate individual model precision, recall, and F1-scores. Detection times (min) along with trial area (m^2) were noted and used to calculate model detection rates (ha/h). Precision, recall, F1-scores, and model detection rate, for each of the four trials were also averaged to obtain the mean and standard deviation of each individual model across the four fields.

3.4   Results

This section will display, describe, and compare the dependent variables (precision, recall, F1-scores, and model detection rates) results for the independent variable "custom trained CNN" (CNNs; RetinaNet, Faster R-CNN, Single Shot Detector, and YOLO v3 each trained on

imagery with 0.4, 0.6, 0.8, and 1.2 cm spatial resolutions). First, the assessment of the

orthomasic derived weed key is reported. Next, the training, model validation, and the initial

model testing are discussed. Finally, the results of model field deployment in four production

fields and resulting precision, recall, and calculated F1-scores are summarized.

3.4.1   Evaluation of orthomosaic derived key

The results of each weed detection experiment were compared to a manually annotated

key that was labeled using the highest spatial resolution (0.4 cm) image sets. An evaluation was

conducted to assess the accuracy of these orthomosaic derived weed labels. Table 3.5

summarizes the true positive, false positive, false negative, and recall results of comparing the

orthomosaic labeled objects to the *in situ* collected ground reference objects. Figure 3.8

illustrates the recall results by weed size. Figure 3.9 illustrates the orthomosaic derived key recall

by specific weed size classes (All sizes, ≥ 2.5 cm, ≥ 5.0 cm, and ≥ 7.5 cm).

Table 3.5: Orthomosaic derived key compared with *in situ* ground observations.

Weed height (cm)	True Positive	False Positive	False Negative	Recall
All	364	12	66	0.83
≥ 2.5	361	15	30	0.92
≥ 5.0	339	37	16	0.96
≥ 7.5	250	126	9	0.97
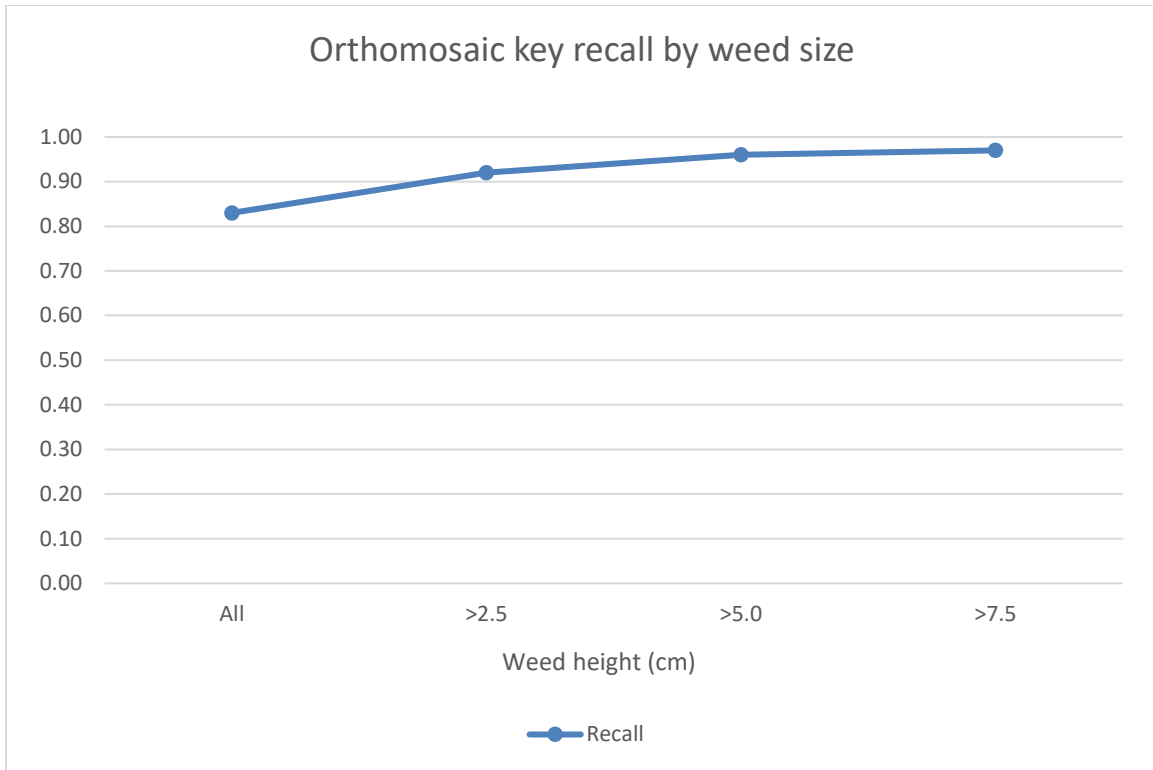
Figure 3.8: Comparison of orthomosaic derived key compared with in situ ground observations.

The evaluated orthomosaic derived key produced a recall of 0.83 for all weeds observed through *in situ* scouting. Most weeds missed by orthomosaic labeling were those that were < 2.5 cm in height. This category primarily consisted of weeds that were emerging and represented by a very small group of or no pixels in orthomosaics at 0.4 cm spatial resolution. Recall increased to 0.92, 0.96, and 0.97 as the weed size threshold was increased to ≥ 2.5, ≥ 5.0, and ≥ 7.5 cm, respectively. Figure 3.9 illustrates the decreasing ability to label weeds < 2.5 cm in height (recall $= 0.08$) in orthomosaics. Weed categories $2.5 > 5.0$ and >5.0 cm in height obtained a recall of 0.86 and 0.93, respectively.
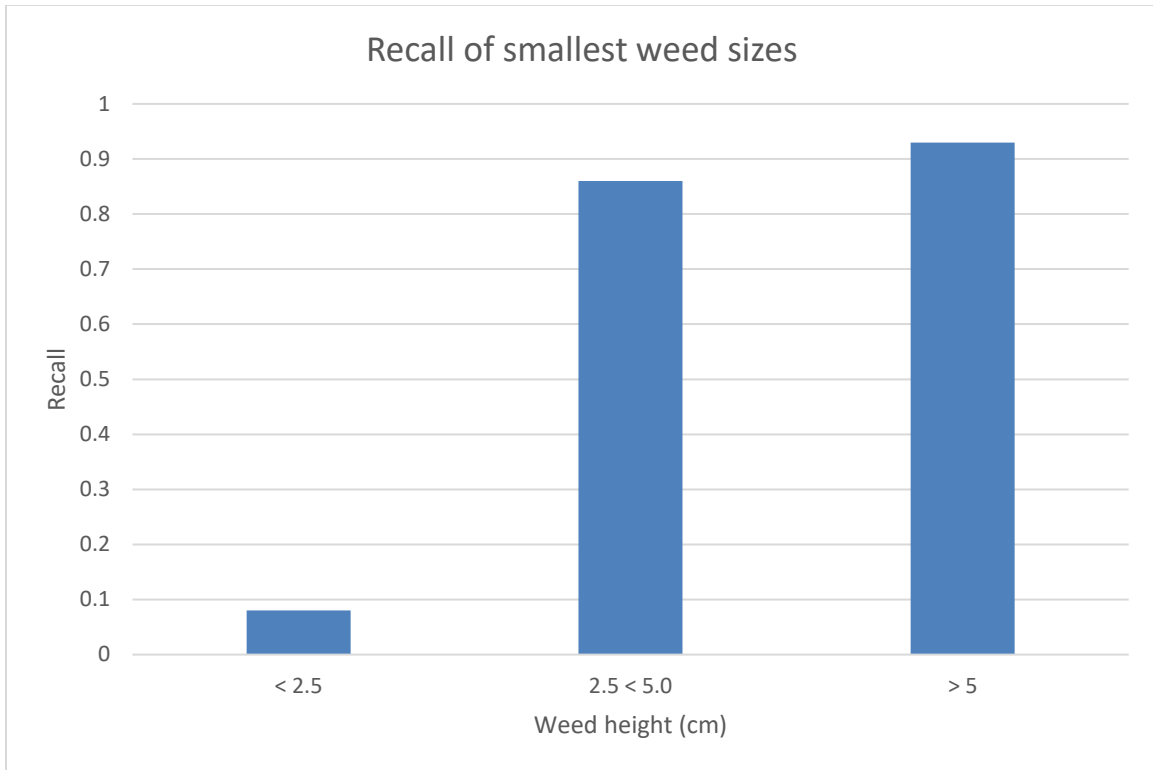
Figure 3.9: Recall of small weeds labeled within the orthomosaic key.

3.4.2   Model validation results

As part of the training process, a subset of imagery was automatically withheld for validation and used to estimate model accuracy while tuning hyperparameters. Once a model finishes training, the results of the validation test after its final epoch are supplied as part of results. Since training was completed with two classes, broadleaf and pigweed, accuracies of both classes are reported. Table 3.6 summarizes the epochs completed, training times, and final validation accuracy results for the 16 trained CNNs. Models RetinaNet at 0.4, 0.8, and 1.2 cm GSD automatically stopped training short of completing 20 epochs because of a lack of improvement. The remaining 13 models completed the maximum 20 epochs.

Training times are illustrated for individual models in Figures 3.10. Faster R-CNN with 0.4 cm spatial resolution imagery had the longest single training time at 401 min, followed by Faster R-CNN trained on 0.6 cm spatial resolution (257 min), YOLO v3 trained on 0.4 cm spatial resolution (225 min), and RetinaNet trained on 0.4 cm spatial resolution (168 min). A consistent decrease in training time is noted with a decrease in imagery spatial resolution for each CNN except for SSD trained on 0.6 cm imagery and YOLO v3 trained on 0.6 cm imagery which deviated from this pattern.

Table 3.6: Training parameters for 16 custom trained CNNs.

Model Number	CNN	Spatial resolution (cm)	Epochs completed (Max 20)	Training Time (min)	Reported Accuracy Pigweed	Reported Accuracy Broadleaf
1	RetinaNet	0.4	16	168	0.60	0.10
2	RetinaNet	0.6	20	101	0.64	0.17
3	RetinaNet	0.8	18	75	0.40	0.00
4	RetinaNet	1.2	19	44	0.24	0.00
5	Faster R-CNN	0.4	20	401	0.80	0.65
6	Faster R-CNN	0.6	20	257	0.84	0.63
7	Faster R-CNN	0.8	20	163	0.82	0.55
8	Faster R-CNN	1.2	20	97	0.78	0.61
9	SSD	0.4	20	108	0.55	0.33
10	SSD	0.6	20	16	0.63	0.35
11	SSD	0.8	20	39	0.47	0.05
12	SSD	1.2	20	21	0.42	0.17
13	YOLO v3	0.4	20	225	0.67	0.44
14	YOLO v3	0.6	20	40	0.70	0.58
15	YOLO v3	0.8	20	83	0.52	0.68
16	YOLO v3	1.2	20	46	0.60	0.53
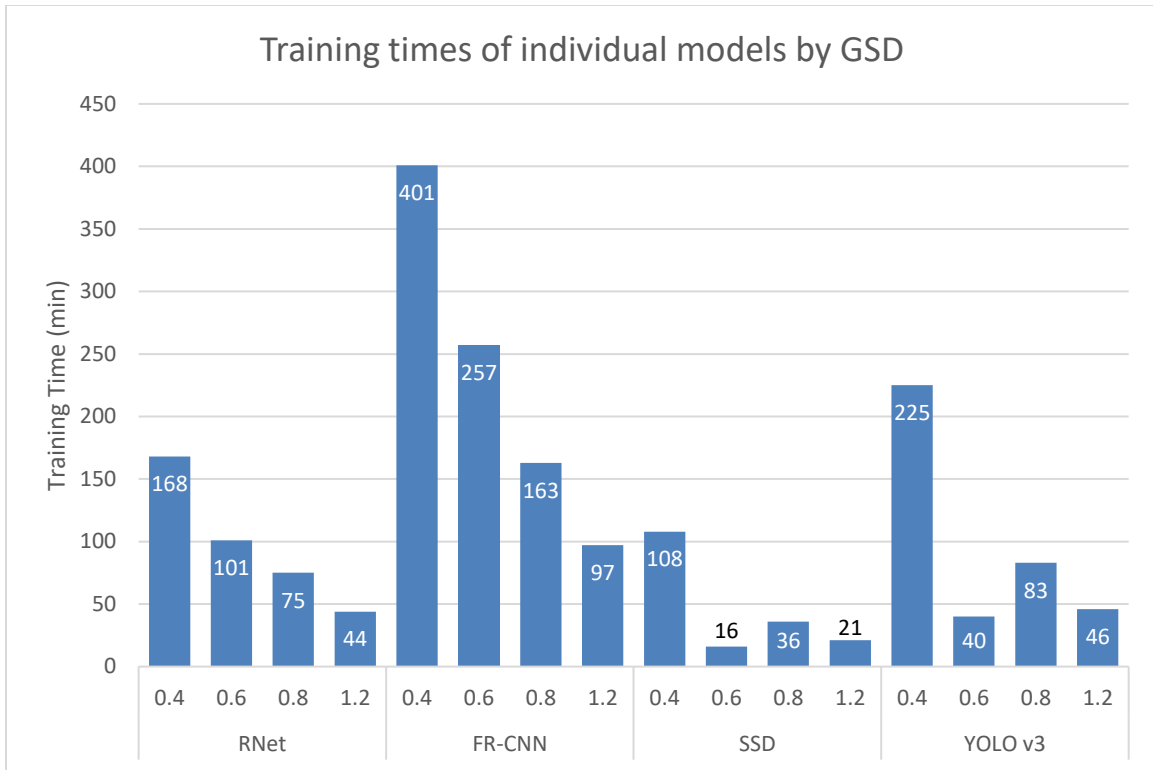
Figure 3.10: Training times of 16 custom trained models. Experimental platform configuration used for testing was the following: OS, Windows 10; GPU NVIDIA Quadro RTX 3000 with Max-Q Design; CPU, Intel® Core™ i7-1065 G7 CPU @ 1.30 GHz with 32 Gb of RAM. Batch size 2 and max epoch 20.

Training losses are the sum of errors at the end of the training session within each epoch. Likewise, validation losses are the sum of errors at the end of validation session within each epoch. Models that fit data better with each subsequent epoch will show parallel reductions in training and validation losses. Each of the CNN architectures tested, calculate these losses differently as represented by their y-axis scale and should not be directly compared numerically. However, the stable decrease in losses over training epochs (measured in batch process) indicates that the data is being modeled with increasing accuracy (Figure 3.11). The loss functions of both Faster R-CNN and YOLO v3 across all GSD training sets showed a steady and stable reduction in both training and validation losses over epochs suggesting good fits to the training data and the parallel decrease in validation losses indicate the ability to generalize. The loss function of

RetinaNet and SSD show irregular spikes in the validation losses particularly at decreased spatial resolutions suggesting poorer fit with the data. This suggests that RetinaNet and SSD was not able to generalize well with this data set compared to the other models evaluated.

The validation results by class are summarized in Table 3.4 and illustrated in Figure 3.12. Except for YOLO v3 at 0.8 GSD, all models had higher pigweed class detection accuracies compared with the broadleaf class. This is likely because the pigweed class was objects (weeds) that were of large enough size and maturity that they were easily and definitively identified as pigweeds. Whereas the more generic broadleaf class were those definitive weeds needing to be targeted but with lower confidence of their specific species, most often due to their size. Lower confidence and poor detection accuracy shown in the broadleaf class illustrates the ambiguous line provided in the imagery between definitive pigweed and definitive weed. Since both classes are targets of herbicide treatment, in subsequent testing, model outputs are combined into a single class and evaluated as such on their detection accuracy.

Considering individual model accuracy results, Faster R-CNN across all resolutions outperformed all other models at pigweed detection. Apart from YOLO v3 trained on 0.8 spatial resolution imagery, Faster R-CNN reported a higher broadleaf detection accuracy than all other models when comparing models trained on equivalent spatial resolution image sets.
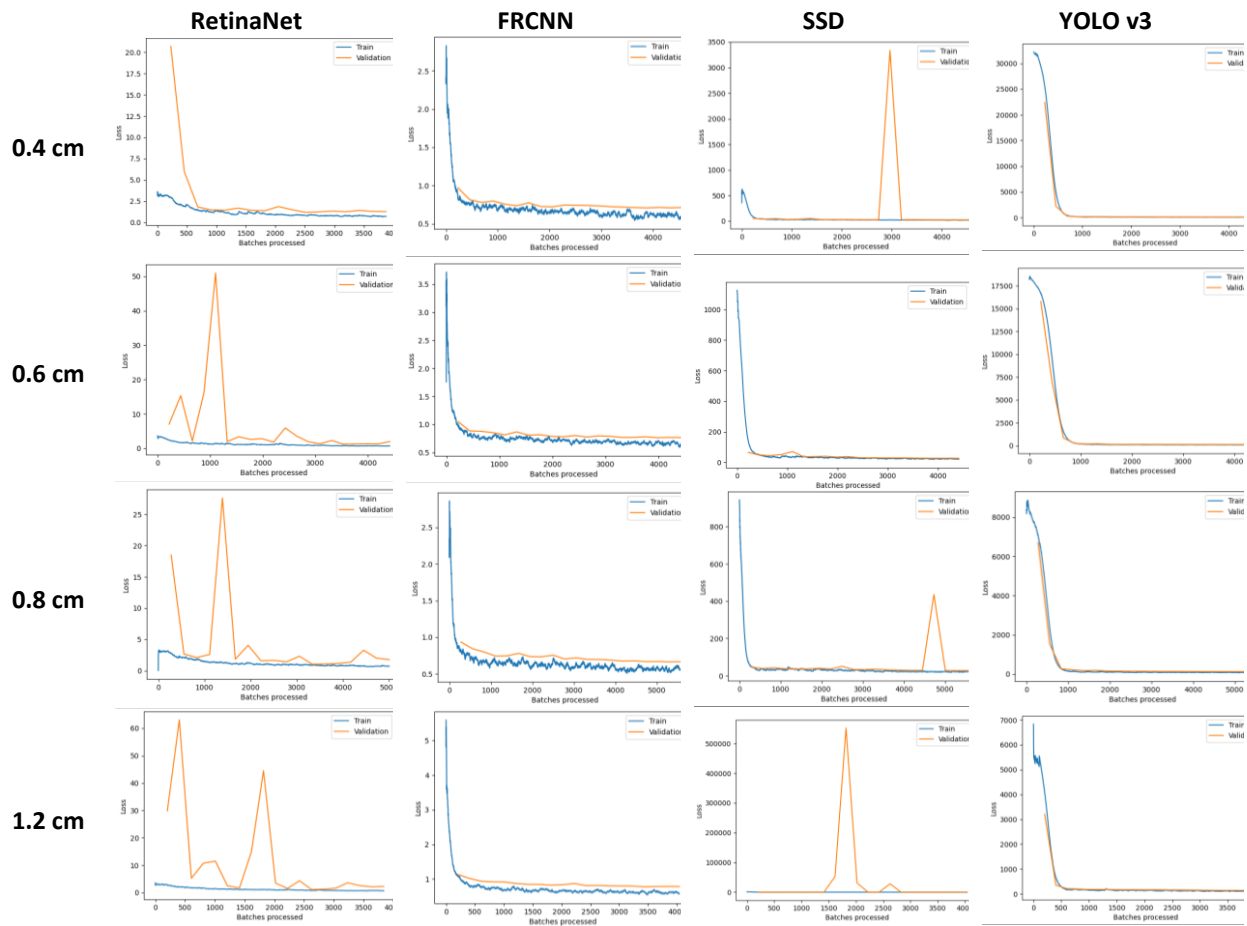
Figure 3.11: Loss functions of 16 models across batch process while training. The loss functions of both Faster R-CNN and YOLO v3 across all GSD training sets showed a steady and stable reduction in both training and validation losses over epochs suggesting good fits to the training data and the parallel decrease in validation losses indicate the ability to generalize. The loss function of RetinaNet and SSD show irregular spikes in the validation losses particularly at decreased spatial resolutions suggesting poorer fit with the data. This suggests that RetinaNet and SSD was not able to generalize well with this data set compared to the other models evaluated.
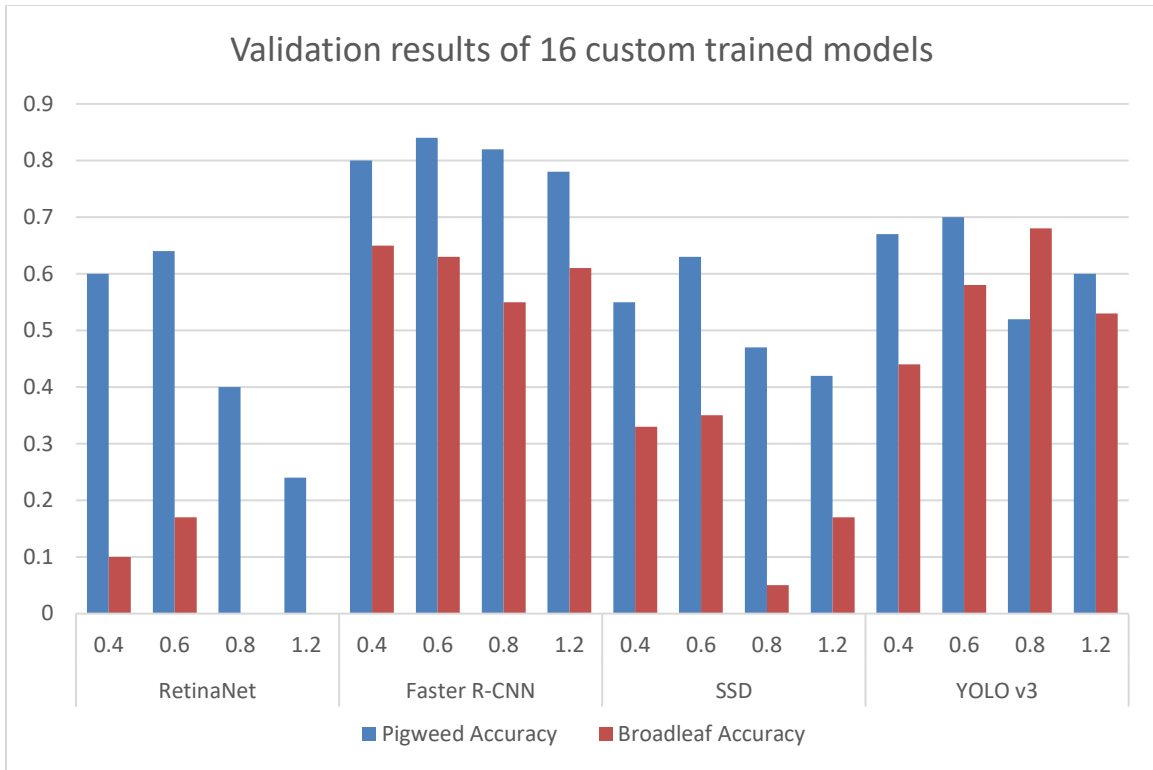
Figure 3.12: Validation results of 16 custom trained models evaluating two classes (pigweed and broadleaf).

3.4.3   Model testing results

Separate from that used for model validation in the training process, ten percent of the training imagery was held for model testing. This imagery was sampled from the same fields and flights as that used for training. Therefore, the results provide some indication of each model's ability to detect weeds that it has not seen but in identical environmental conditions that were presented in training. This section summarizes the results of model testing using this ten percent of training imagery withheld for testing. Detection times, true positive, false positive, false negative, precision, recall, and accuracy (F1-score) from these tests are displayed in Table 3.7.

58

Table 3.7: Model testing results for 16 custom trained models (0.05 detection threshold)

Model Number	CNN	Resolution (cm GSD)	Detection Time (min)	True Positives	False Positives	False Negatives	Precision	Recall	Accuracy (F1-Score)
1	RetinaNet	0.4	11.50	582	17992	229	0.03	0.72	0.06
2	RetinaNet	0.6	6.50	499	7397	313	0.06	0.62	0.11
3	RetinaNet	0.8	2.75	295	4696	517	0.06	0.36	0.10
4	RetinaNet	1.2	2.50	146	4057	666	0.03	0.18	0.06
5	Faster R-CNN	0.4	18.00	762	188	34	0.80	0.96	0.87
6	Faster R-CNN	0.6	9.00	678	224	118	0.75	0.85	0.80
7	Faster R-CNN	0.8	6.00	666	309	130	0.68	0.84	0.75
8	Faster R-CNN	1.2	3.50	587	314	210	0.65	0.74	0.69
9	SSD	0.4	5.75	331	257	331	0.56	0.50	0.53
10	SSD	0.6	2.00	225	354	439	0.37	0.34	0.35
11	SSD	0.8	2.00	9	109	655	0.08	0.01	0.20
12	SSD	1.2	1.50	101	594	563	0.15	0.15	0.15
13	YOLO v3	0.4	9.45	484	450	163	0.92	0.86	0.89
14	YOLO v3	0.6	4.45	584	59	63	0.86	0.82	0.84
15	YOLO v3	0.8	3.00	459	149	188	0.75	0.68	0.71
16	YOLO v3	1.2	1.75	285	349	363	0.45	0.42	0.43

Precision

True positive and false positive detections were used to calculate precision results. Precision metrics grouped relatively close by CNN (Figure 3.13). Comparing models trained on imagery of the same spatial resolution, YOLO v3 reported the highest precision results (0.92, 0.86, and 0.75) for spatial resolutions 0.4, 0.6, and 0.8, respectively. Faster R-CNN reported the second highest and most tightly grouped precision results across spatial resolutions (0.80, 0.75, 0.68, and 0.65) for spatial resolutions 0.4, 0.6, 0.8, and 1.2 cm. The top three individual model precision results were YOLO v3 trained on 0.4 cm spatial resolution imagery (0.92), YOLO v3 trained on 0.6 cm spatial resolution imagery (0.86), and Faster R-CNN trained on 0.4 cm spatial resolution imagery (0.80). RetinaNet precision results were consistently the poorest with no model's precision results exceeding 0.06.

Figure 3.13: Precision of model testing for 16 models (four CNNs trained on imagery of spatial resolutions 0.4, 0.6, 0.8, and 1.2 cm).

Recall

True positive and true negative detections were used to calculate recall results. Individual model recall results ranked consistently by CNN when comparing respective resolutions with Faster R-CNN performing the best, followed by YOLO v3, RetinaNet, and SSD (Figure 3.14). Models trained on the highest spatial resolution imagery (0.4 cm GSD) produced a recall of 0.96, 0.86, 0.72, and 0.50 for Faster R-CNN, YOLO v3, RetinaNet, and SSD, respectively. Subsequent resolutions ranked CNNs identically. The top three recall results were produced by Faster R-CNN at 0.4 cm GSD, YOLO v3 at 0.4 cm GSD, and Faster R-CNN at 0.6 cm GSD (0.96, 0.86, and 0.85 respectively).

Figure 3.14: Recall of model testing for 16 models (four CNNs trained on imagery of spatial resolutions 0.4, 0.6, 0.8, and 1.2 cm).

F1-Scores

F1-scores were calculated using individual model precision and recall results. Individual model F1-scores closely mirrored individual precision results (Figure 3.15). YOLO v3 trained on 0.4 cm spatial resolution imagery produced the highest F1 measure (F1 = 0.89), followed by Faster R-CNN trained on 0.4 cm spatial resolution imagery (F1 = 0.87), YOLO v3 trained on 0.6 cm spatial resolution imagery (F1 = 0.84), and then Faster R-CNN trained on 0.6 cm spatial resolution imagery (F1 = 0.80). The remaining models all produced F1 measures below 0.76. RetinaNet consistently produced the lowest F1 measure with no model exceeding 0.11.
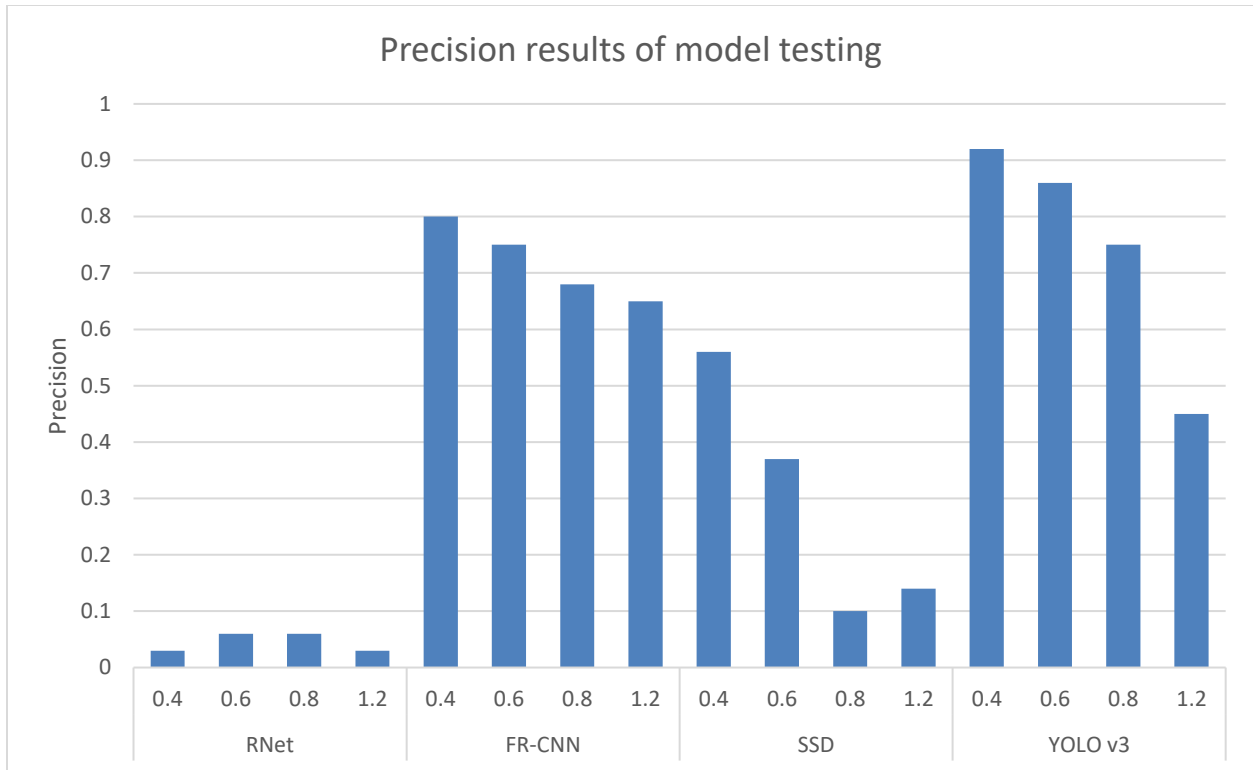
Figure 3.15: F1-scores of model testing for 16 models (four CNNs trained on imagery of spatial resolutions 0.4, 0.6, 0.8, and 1.2 cm).

Detection times

As shown in Figure 3.16, model detection times (min) are reported instead of detection rates (ha/h) for each CNN. Since, training chips were an aggregation of multiple flights, it was not feasible to estimate imaged area (ha) to calculate detection rates. For detection times within each CNN, times consistently increased with increasing spatial resolution. The most notable jump in detection times of CNNs were between models trained on 0.4 and 0.6 cm spatial resolution. Comparing models of the same spatial resolution, Faster R-CNN's detection times were longer than any other CNN followed by RetinaNet, YOLO v3, and SSD.
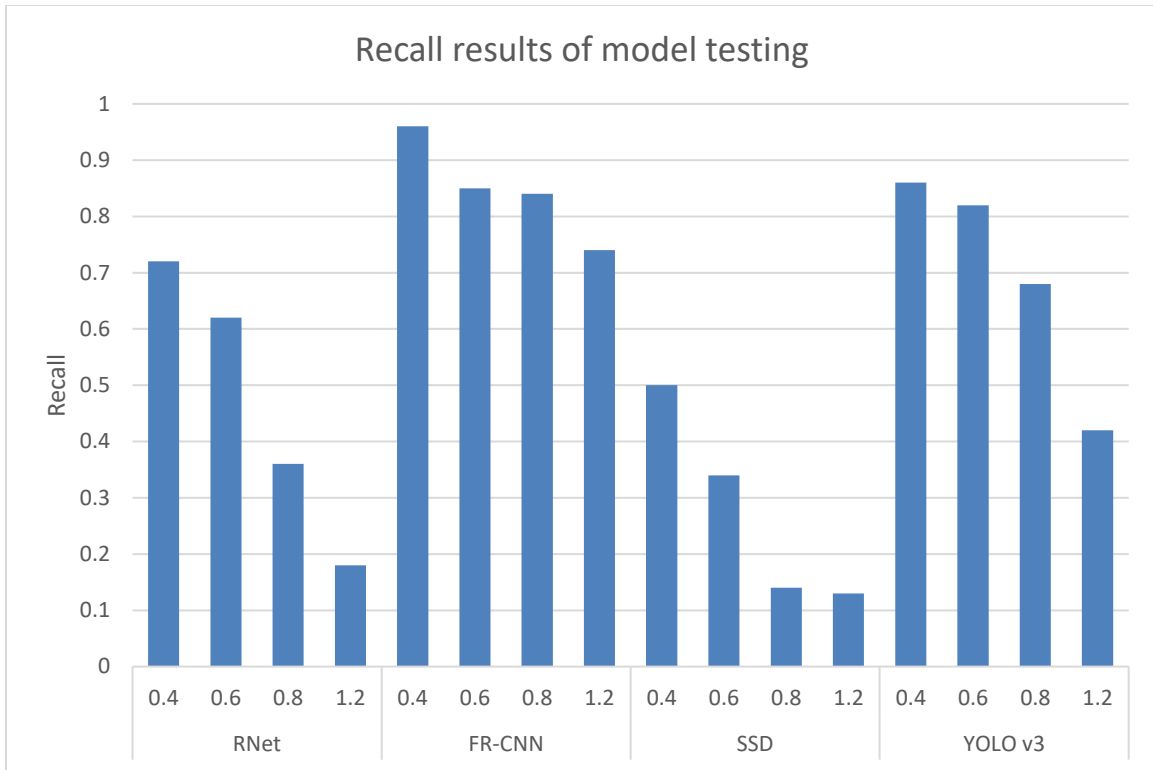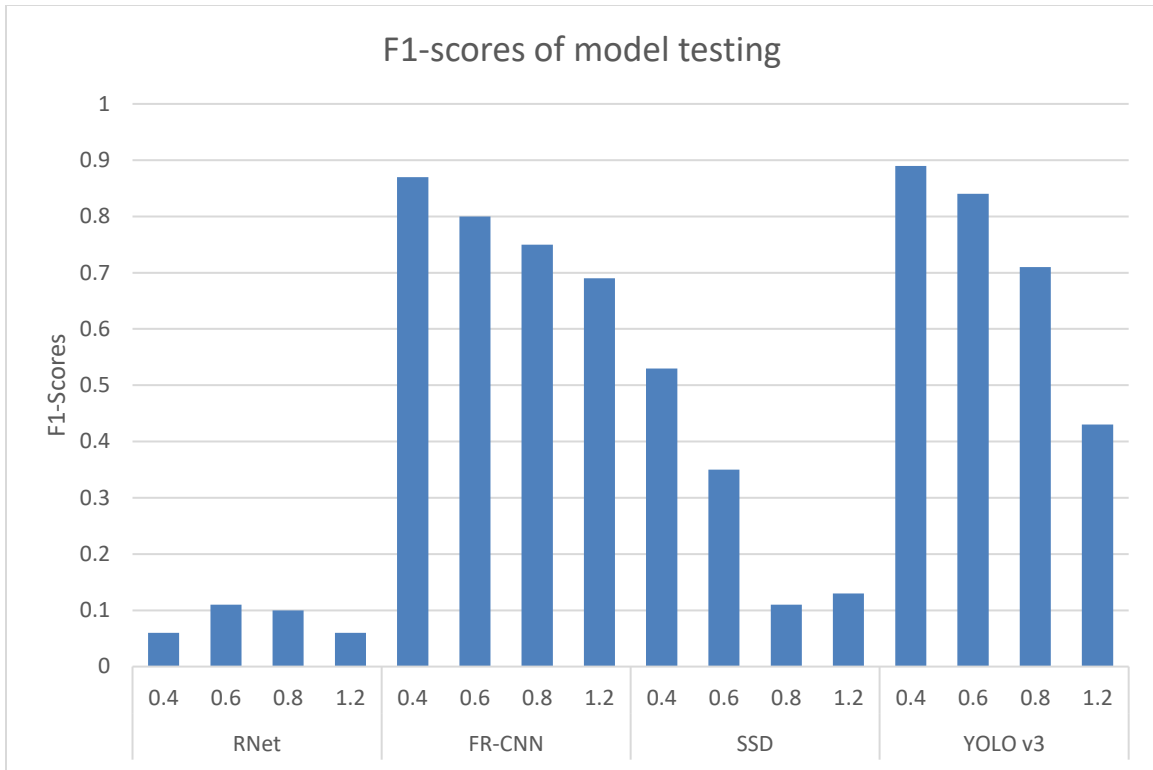
62
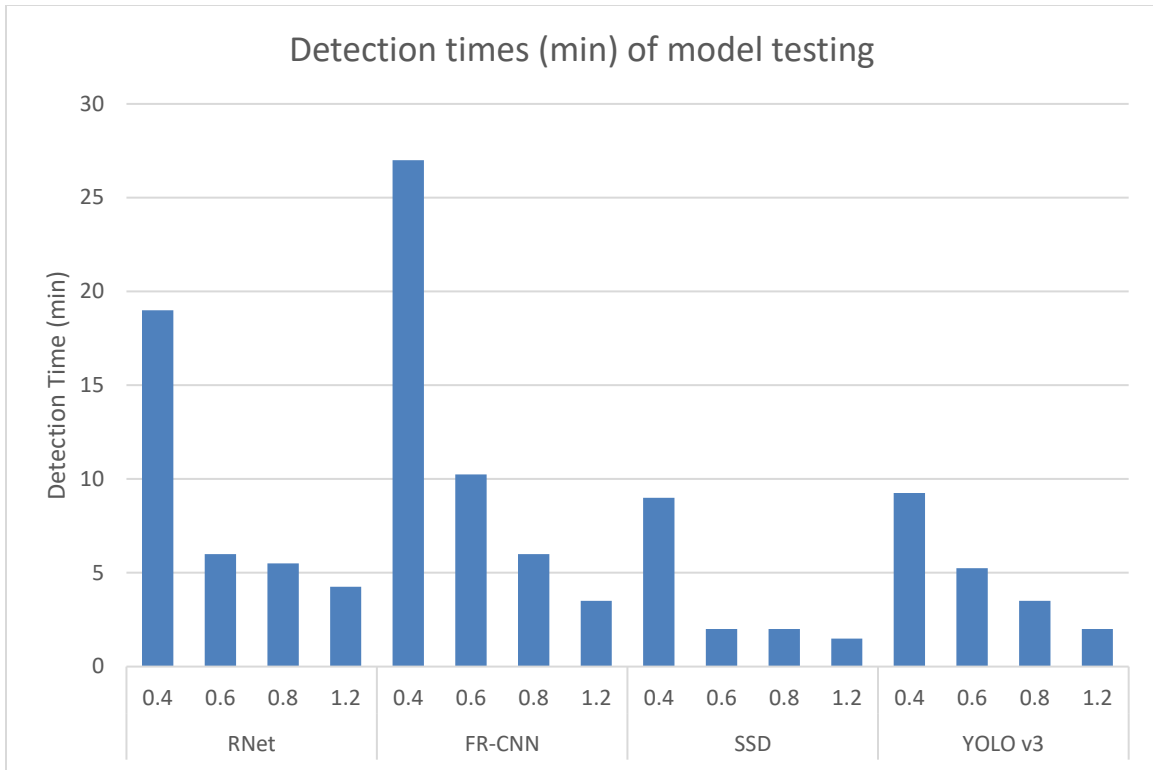
Figure 3.16: Detection times (min) of model testing for 16 models (four CNNs trained on imagery of spatial resolutions 0.4, 0.6, 0.8, and 1.2 cm).

3.4.4 Model deployment results

The 16 custom trained models were deployed in four different production fields. The individual detection results of each field deployment and aggregate detection are discussed in this section. Field variability of the four deployment fields including crop size, weed size, soybean row development, soil, plant and weed color contrast. Additionally, some fields contained miscellaneous weeds that were not included in training imagery. This variability between individual fields is described as part of detection results to provide indications of expected model performance in other similar fields and future deployments. Analysis of aggregate results is completed to evaluate models' ability to generalize across the deployment field conditions.

3.4.5 Individual field results from for field deployments

Soybean plants in field one where in late vegetative stages (six to seven trifoliate leaves) and large enough that they had consolidated into continuous rows (Figure 3.17). Pigweeds were numerous with a weed density of 0.96 weeds/m^2. Soil color in imagery was bright due to low soil moisture and contrasted well with plants. Pigweed color on the day of collection appeared a brighter green than soybean plants. Because of imagery collection time occurring near local solar noon, shadows were minimal. Additionally, it was noted that there were evenly distributed examples of *Sorghum halepense* (Johnson grass) and sparce examples of *Ipomoea* (morning glory) present, neither were represented in training imagery. The *Sorghum halepense* were appropriately ignored by the model since no grass was targeted, while the *Ipomoea* (morning glory) were properly grouped into the broadleaf weed category. Once detections were made, classes were grouped and evaluated as single class.
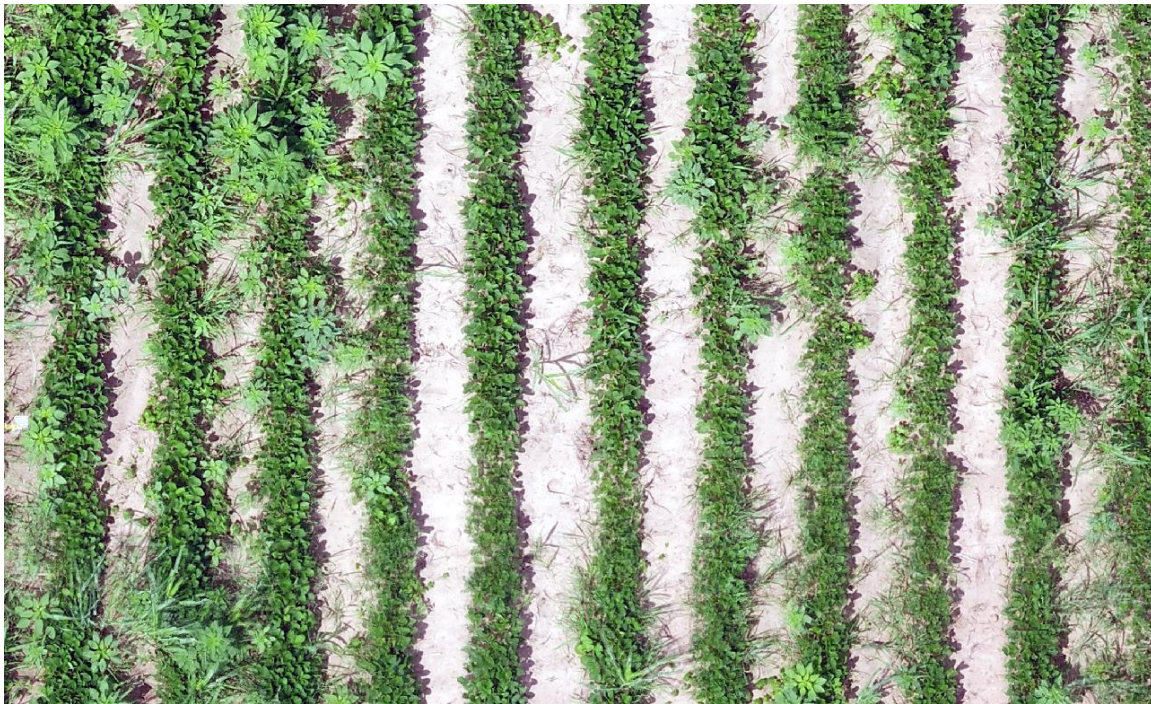
Figure 3.17: Ten-row example image from field one. Soybean plants have merged into continuous rows, weeds are relatively large, and grass weeds are present.

Field one precision, recall, F1-scores, and detection rates (ha/h) are displayed in Table

3.8. The highest precision results for field one deployment were achieved by Faster R-CNN

trained on 0.4 cm spatial resolution imagery (precision = 0.89), followed by Faster R-CNN

trained on 0.6 (precision = 0.75) and 0.8 (precision = 0.69) cm spatial resolution. The top three

recall results for the individual models were Faster R-CNN trained on 0.4 spatial resolution

imagery (recall = 0.99), YOLO v3 trained on 0.4 spatial resolution imagery (recall = 0.90) and

Faster R-CNN trained on 0.6 spatial resolution imagery (0.88). The top three individual model

F1-score ranking were identical to precision rankings, Faster R-CNN trained on 0.4 cm spatial

resolution (F1-score = 0.94), followed by Faster R-CNN trained on 0.6 (F1-score = 0.82) and 0.8

(F1-score= 0.76) cm spatial resolution. Detection rates (ha/h) increased as spatial resolution

coarsened across all CNNs. The fastest detection rates were achieved by YOLO v3 trained on 1.2

cm spatial resolution at 8.82 ha/h, followed by SSD trained on 1.2 cm spatial resolution imagery

at 8.01 ha/h, and SSD trained on 0.6 cm spatial resolution imagery at 6.78 ha/h. RetinaNet

trained on 0.4 cm spatial resolution imagery produced the slowest detection rate of 0.79 ha/h,

followed by Faster R-CNN trained on 0.4 spatial resolution imagery (detection rate = 1.66).

Table 3.8: Field deployment results for 16 custom trained models on field one (IoU 0.5). Highest performance metrics are bolded.

Model Number	Model	Resolution (cm GSD)	Detection rate (ha/h)	Precision	Recall	F1-Score
1	RetinaNet	0.4	0.79	0.02	0.70	0.03
2	RetinaNet	0.6	2.32	0.04	0.61	0.08
3	RetinaNet	0.8	3.67	0.04	0.36	0.07
4	RetinaNet	1.2	5.51	0.02	0.14	0.04
5	Faster R-CNN	0.4	1.66	**0.89**	**0.99**	**0.94**
6	Faster R-CNN	0.6	2.84	0.75	0.88	0.82
7	Faster R-CNN	0.8	3.53	0.69	0.86	0.76
8	Faster R-CNN	1.2	6.30	0.57	0.73	0.64
9	SSD	0.4	3.53	0.40	0.26	0.32
10	SSD	0.6	6.78	0.33	0.28	0.30
11	SSD	0.8	6.30	0.12	0.15	0.14
12	SSD	1.2	8.01	0.09	0.14	0.12
13	YOLO v3	0.4	2.32	0.62	0.90	0.73
14	YOLO v3	0.6	4.41	0.63	0.67	0.55
15	YOLO v3	0.8	5.88	0.48	0.57	0.52
16	YOLO v3	1.2	8.82	0.25	0.39	0.31

Soybean plants in field two imagery were more mature than field one and were at R1 growth stage (beginning flowering) (Figure 3.18). More mature plants produced more uniform rows and narrower views of furrows. Pigweeds were of similar size to field one but were sparce by comparison (weed density = 0.42 weeds/m^2). There were *Solanum carolinense* (horse nettle) broadleaf weeds present in field two imagery. Pigweed color and leaf size closely resembled soybean plants. Soil color in field two imagery was blotchy and darker due to higher moisture content. Crop residue left over from previous seasons harvest was present as well as some soybean leaves that had dropped triggering some false positives and reducing precision.

Figure 3.18: Ten-row example image from field two. Soybean plants are more mature, pigweeds sparce, and soil is of varying color and contains residue.

Field two precision, recall, F1-scores, and detection rates (ha/h) are displayed in Table 3.9. Precision results dropped steeply on field two imagery across all CNNs with the highest precision results peaking at 0.51 achieved by YOLO v3 trained on 0.4 cm spatial resolution, followed by SSD trained on 0.4 (precision = 0.47) and Faster R-CNN trained on 0.4 cm spatial resolution (precision = 0.40). The top three recall results for the individual models were Faster R-CNN trained on 0.4, 0.6 and 0.8 spatial resolution imagery (recall = 0.98, 0.88, 0.81, respectively). Like in field one, the top three individual model F1-score ranking were identical to precision rankings, YOLO v3 trained on 0.4 cm spatial resolution (F1-score = 0.59), followed by Faster R-CNN trained on 0.4 spatial resolution (F1-score = 0.56) and SSD trained on 0.8 cm spatial resolution (F1-score= 0.42). Detection rates (ha/h) increased as spatial resolution coarsened across all CNNs with only a couple exceptions (SSD and YOLO v3 trained on 0.6 and

0.8 cm spatial resolutions). Fastest detection rates were achieved by YOLO v3, and SSD trained on 1.2 cm spatial resolution (both 3.03 ha/h), Followed by Faster R-CNN trained on 1.2 cm spatial resolution imagery (detection rate = 2.48 ha/h). RetinaNet trained on 0.4, and 0.6 spatial resolutions and Faster R-CNN at 0.4 GSD produced the slowest three detection rates of 0.36, 0.62, and 0.62 ha/h, respectively.

Table 3.9: Field deployment results for 16 custom trained models on trial two (IoU 0.5). Highest performance metrics are bolded.

Model Number	Model	Resolution (cm GSD)	Detection rate (ha/h)	Precision	Recall	F1-Score
1	RetinaNet	0.4	0.36	0.00	0.64	0.01
2	RetinaNet	0.6	0.62	0.01	0.50	0.01
3	RetinaNet	0.8	0.68	0.00	0.36	0.01
4	RetinaNet	1.2	1.30	0.00	0.19	0.01
5	Faster R-CNN	0.4	0.62	0.40	**0.98**	0.56
6	Faster R-CNN	0.6	0.64	0.12	0.88	0.21
7	Faster R-CNN	0.8	1.30	0.13	0.81	0.22
8	Faster R-CNN	1.2	2.48	0.07	0.60	0.13
9	SSD	0.4	0.85	0.47	0.38	0.42
10	SSD	0.6	1.95	0.08	0.23	0.07
11	SSD	0.8	1.30	0.05	0.15	0.07
12	SSD	1.2	3.03	0.02	0.08	0.03
13	YOLO v3	0.4	0.80	**0.51**	0.70	**0.59**
14	YOLO v3	0.6	1.24	0.21	0.52	0.30
15	YOLO v3	0.8	1.09	0.17	0.42	0.25
16	YOLO v3	1.2	3.03	0.09	0.25	0.13

Soybean plants in field three where in late vegetative stages (five to six trifoliate leaves) and were large enough that they had consolidated into continuous rows (Figure 3.19). Weed

68

density in field three was 0.64 weeds/m^2. Soil color in imagery was bright due to low soil moisture, smooth, and contrasted well with plants. Pigweed color on the day of collection appeared a brighter green than soybean plants. There was some minimal shading on one side of the soybean rows. No additional weed species were noted by *in situ* field observation, only pigweed.
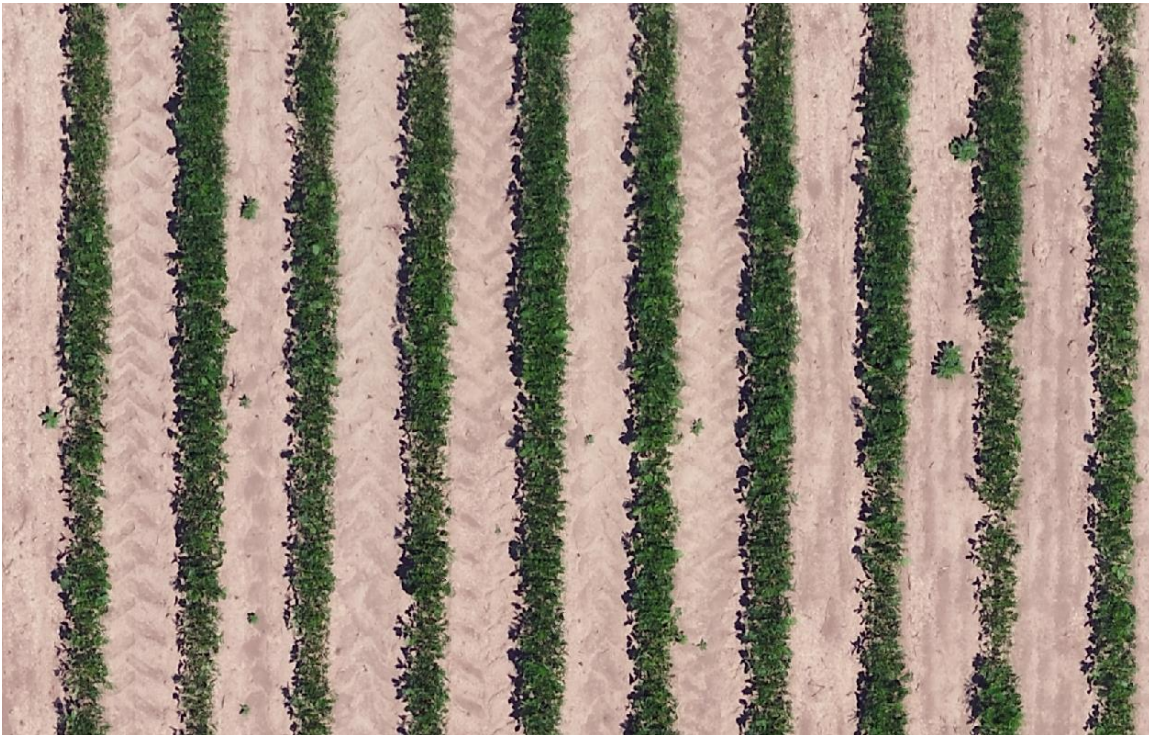
Figure 3.19: Ten-row example image from field three. Soybean plants are consolidated into continuous rows, pigweeds are small and evenly distributed. Soil is bright, smooth and contrasts well with plants.

Field three precision, recall, F1-scores, and detection rates (ha/h) are displayed in Table 3.10. The highest precision results for field two deployment were achieved by Faster R-CNN trained on 0.4 cm spatial resolution (precision = 0.97), followed by YOLO v3 trained on 0.4 (precision = 0.92) and 0.6 (precision = 0.86) cm spatial resolution. The top three recall results for the individual models were Faster R-CNN trained on 0.4, 0.6, and 0.8 cm spatial resolution

imagery (recall = 0.98, 0.91, and 0.89, respectively). All remaining model recalls were below 0.80. Top three individual model F1-score rankings were identical to recall rankings, Faster R-CNN trained on 0.4, 0.6, and 0.8 cm spatial resolutions produced F1-scores of 0.98, 0.87, 0.84, respectively. Detection rates (ha/h) increased as spatial resolution coarsened across all CNNs with one exception (SSD trained on 0.6 and 0.8 cm spatial resolution imagery). The fastest detection rates were achieved by SSD trained on 1.2 cm spatial resolution at 11.72 ha/h, followed by YOLO v3 trained on 1.2 cm spatial resolution imagery at 10.75 ha/h, and then SSD trained on 0.6 cm spatial resolution imagery at 8.06 ha/h. RetinaNet trained on 0.4 cm spatial resolution imagery produced the slowest detection rate of 1.43 ha/h, followed by Faster R-CNN trained on 0.4 spatial resolution imagery (detection rate = 1.65).

Table 3.10: Field deployment results for 16 custom trained models on trial three (IoU 0.5). Highest performance metrics are bolded.

Model Number	Model	Resolution (cm GSD)	Detection rate (ha/h)	Precision	Recall	F1-Score
1	RetinaNet	0.4	1.43	0.02	0.61	0.03
2	RetinaNet	0.6	3.39	0.04	0.64	0.08
3	RetinaNet	0.8	4.16	0.02	0.27	0.03
4	RetinaNet	1.2	6.14	0.01	0.09	0.02
5	Faster R-CNN	0.4	1.65	**0.97**	**0.98**	**0.98**
6	Faster R-CNN	0.6	3.07	0.84	0.91	0.87
7	Faster R-CNN	0.8	4.30	0.80	0.89	0.84
8	Faster R-CNN	1.2	7.16	0.80	0.77	0.78
9	SSD	0.4	5.37	0.43	0.17	0.25
10	SSD	0.6	8.06	0.36	0.25	0.29
11	SSD	0.8	7.16	0.08	0.07	0.07
12	SSD	1.2	11.72	0.14	0.09	0.11
13	YOLO v3	0.4	2.93	0.92	0.63	0.75
14	YOLO v3	0.6	5.16	0.86	0.66	0.75
15	YOLO v3	0.8	6.79	0.84	0.60	0.70
16	YOLO v3	1.2	10.75	0.54	0.43	0.48

Soybean plants in field four were late vegetative (six to seven trifoliate leaves); however, they were slightly stunted due to drought conditions and located in a part of the field that was not irrigated. Smaller plant size caused rows to be broken and of irregular width (Figure 3.20). Soybean plants and weeds appeared darker and close in color. The vast majority of weeds in field four were young pigweeds with a field weed density of 0.24 weeds/m^2. There were also examples of scattered *Solanum carolinense* (horsenettle) weeds. The soil color in the imagery was bright due to low soil moisture and contrasted well with plants.

Figure 3.20: Ten-row example image from field four. Soybean plants are intermittently consolidated into continuous rows with some breaks, pigweeds are small and evenly distributed. Soil is bright, smooth and contrasts well with plants.

Field four precision, recall, F1-scores, and detection rates (ha/h) are displayed in Table 3.11. The highest precision results for field four deployment were achieved by Faster R-CNN trained on 0.4, 0.6, and 1.2 cm spatial resolution (precision = 0.84, 0.76, and 0.66). The top three recall results for the individual models were Faster R-CNN trained on 0.4, 0.8, and 0.6 cm spatial resolution (recall = 0.90, 0.89, and 0.88, respectively). The top three individual model F1-score ranking were Faster R-CNN trained on 0.4, 0.6, and 0.8 cm spatial resolutions (F1-score = 0.87, 0.82, 0.75). Detection rates (ha/h) increased as spatial resolution coarsened across all CNNs. The fastest detection rates were achieved by SSD trained on 1.2 cm spatial resolution at 4.18 ha/h, followed by RetinaNet and YOLO v3 trained on 1.2 cm spatial resolution imagery (detection rate = 3.42 and 3.42, respectively). RetinaNet, Faster R-CNN, and YOLO v3 trained on 0.4 cm

spatial resolution imagery produced the slowest detection rate of 0.52, 0.67, and 0.82 ha/h,

respectively.

Table 3.11: Field deployment results for 16 custom trained models on trial four (IoU 0.5). Highest performance metrics are bolded.

Model Number	Model	Resolution (cm GSD)	Detection rate (ha/h)	Precision	Recall	F1-Score
1	RetinaNet	0.4	0.52	0.00	0.75	0.01
2	RetinaNet	0.6	1.11	0.01	0.70	0.01
3	RetinaNet	0.8	1.25	0.00	0.25	0.01
4	RetinaNet	1.2	3.42	0.00	0.00	0.00
5	Faster R-CNN	0.4	0.67	**0.84**	**0.90**	**0.87**
6	Faster R-CNN	0.6	1.18	0.76	0.88	0.82
7	Faster R-CNN	0.8	1.57	0.64	0.89	0.75
8	Faster R-CNN	1.2	2.51	0.66	0.74	0.74
9	SSD	0.4	1.88	0.03	0.02	0.02
10	SSD	0.6	2.51	0.03	0.12	0.04
11	SSD	0.8	2.89	0.00	0.01	0.00
12	SSD	1.2	4.18	0.01	0.02	0.02
13	YOLO v3	0.4	0.82	0.16	0.47	0.23
14	YOLO v3	0.6	1.79	0.59	0.59	0.59
15	YOLO v3	0.8	2.51	0.29	0.45	0.35
16	YOLO v3	1.2	3.42	0.17	0.39	0.24

Aggregate results of four field deployments

Table 3.12 summarizes the results of aggregating true positive, false positive, and false

negative detection results over the four field deployments containing a sum of 7,800 weeds.

Analysis of aggregate results was completed to evaluate models' ability to generalize across the

variability represented by the four deployment field conditions described in the previous section.

Hypothesis testing was conducting using these aggregated results.

Table 3.12: Aggregate field test results for 16 custom trained models (IoU 0.5). Highest performance metrics are bolded.

Model Number	Model	Resolution (cm GSD)	Detection time (min)	Precision	Recall	F1-Score
1	RetinaNet	0.4	18.00	0.01	0.66	0.02
2	RetinaNet	0.6	8.50	0.03	0.62	0.05
3	RetinaNet	0.8	7.50	0.02	0.32	0.03
4	RetinaNet	1.2	2.75	0.01	0.12	0.02
5	Faster R-CNN	0.4	14.00	**0.86**	**0.98**	**0.91**
6	Faster R-CNN	0.6	8.00	0.59	0.89	0.71
7	Faster R-CNN	0.8	6.00	0.58	0.87	0.69
8	Faster R-CNN	1.2	3.75	0.47	0.74	0.57
9	SSD	0.4	5.00	0.38	0.21	0.27
10	SSD	0.6	3.75	0.23	0.25	0.24
11	SSD	0.8	3.25	0.08	0.11	0.09
12	SSD	1.2	**2.25**	0.08	0.11	0.09
13	YOLO v3	0.4	11.50	0.63	0.74	0.68
14	YOLO v3	0.6	5.25	0.64	0.65	0.65
15	YOLO v3	0.8	3.75	0.53	0.57	0.55
16	YOLO v3	1.2	2.75	0.30	0.40	0.34

Precision

Null hypothesis one stated that none of 16 custom trained open-source models (four CNNs trained on imagery of four spatial resolutions) will achieve ≥ 0.80 precision results at *Palmer amaranth* detection. Precision results for CNNs trained on imagery of four spatial resolutions are summarized in Table 3.13 and Figure 3.21. Model Faster R-CNN trained on 0.4

74

cm spatial resolution imagery produced a precision result of 0.86; therefore, null hypothesis one was rejected. Individual model precision values trended lower for each CNN as the imagery they were trained on decreased in spatial resolution. Faster R-CNN and YOLO v3 produced the only precision results that exceeded 0.50. RetinaNet consistently produced the lowest precision results with a maximum of 0.03 while SSD models maxed precision scores at 0.38.

Table 3.13: Precision results of aggregated field deployments organized by CNNs and training imagery spatial resolution (cm).

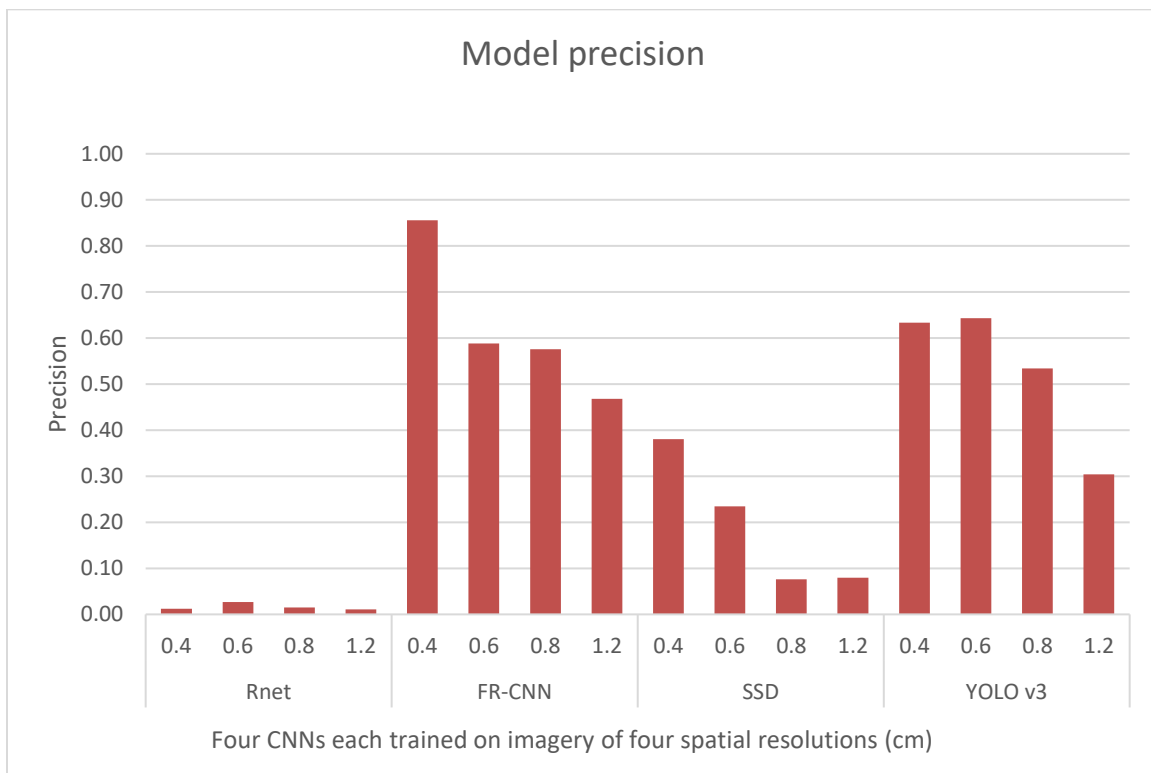|	RetinaNet	Faster R-CNN	SSD	YOLO v3
0.4 cm	0.01	0.86	0.38	0.63
0.6 cm	0.03	0.59	0.23	0.64
0.8 cm	0.02	0.58	0.08	0.53
1.2 cm	0.01	0.47	0.08	0.30

Figure 3.21: Precision results of 16 models evaluated on imagery from four production fields with 7,800 weeds.

Recall

Null hypothesis two stated that none of 16 custom trained open-source models (four CNNs trained on imagery of four spatial resolutions) will achieve ≥ 0.95 recall results at *Palmer amaranth* detection in soybean field imagery. Recall results from CNNs trained on imagery of increasing spatial resolution are summarized in Tables 3.14 and Figures 3.22. All Faster R-CNN models produced recall results higher than 0.70, along with YOLO v3 trained on 0.4 cm spatial resolution imagery. Faster R-CNN models trained on finer spatial resolution imagery 0.4, 0.6, and 0.8 each produced recall results higher than 0.80. Faster R-CNN trained on 0.4 cm spatial resolution produced the highest single model recall result of 0.98; therefore, null hypothesis two was rejected. Individual model recall values trended lower for each CNN as the imagery they were trained on decreases in spatial resolution. When comparing CNNs trained on equivalent spatial resolution SSD produced the poorest recall.

Table 3.14: Recall results of aggregated field deployments organized by CNN and training imagery spatial resolutions (cm).

	RetinaNet	Faster R-CNN	SSD	YOLO v3
0.4 cm	0.66	0.98	0.21	0.74
0.6 cm	0.62	0.89	0.25	0.65
0.8 cm	0.32	0.87	0.11	0.57
1.2 cm	0.12	0.74	0.11	0.40

Figure 3.22: Recall results of 16 models evaluated on imagery from four production fields with 7,800 weeds.

F1-Score

Null hypothesis three stated that none of 16 custom trained open-source models (four CNNs trained on imagery of four spatial resolutions) will achieve ≥ 0.90 F1-score results at *Palmer amaranth* detection in soybean field imagery. F1-scores for four CNNs trained on imagery of 0.4, 0.6, 0.8, and 1.2 cm spatial resolution are summarized in Table 3.15 and Figure 3.23. Faster R-CNN trained on imagery of 0.4 cm spatial resolution produced an F1-score of 0.91; therefore, null hypothesis three was rejected. No other models exceeded a F1-score of 0.71. Except for RetinaNet models, F1-scores increased for CNNs with increasing spatial resolution training imagery.

Table 3.15: F1-score of aggregated field deployments organized by CNNs and training imagery spatial resolution (cm).

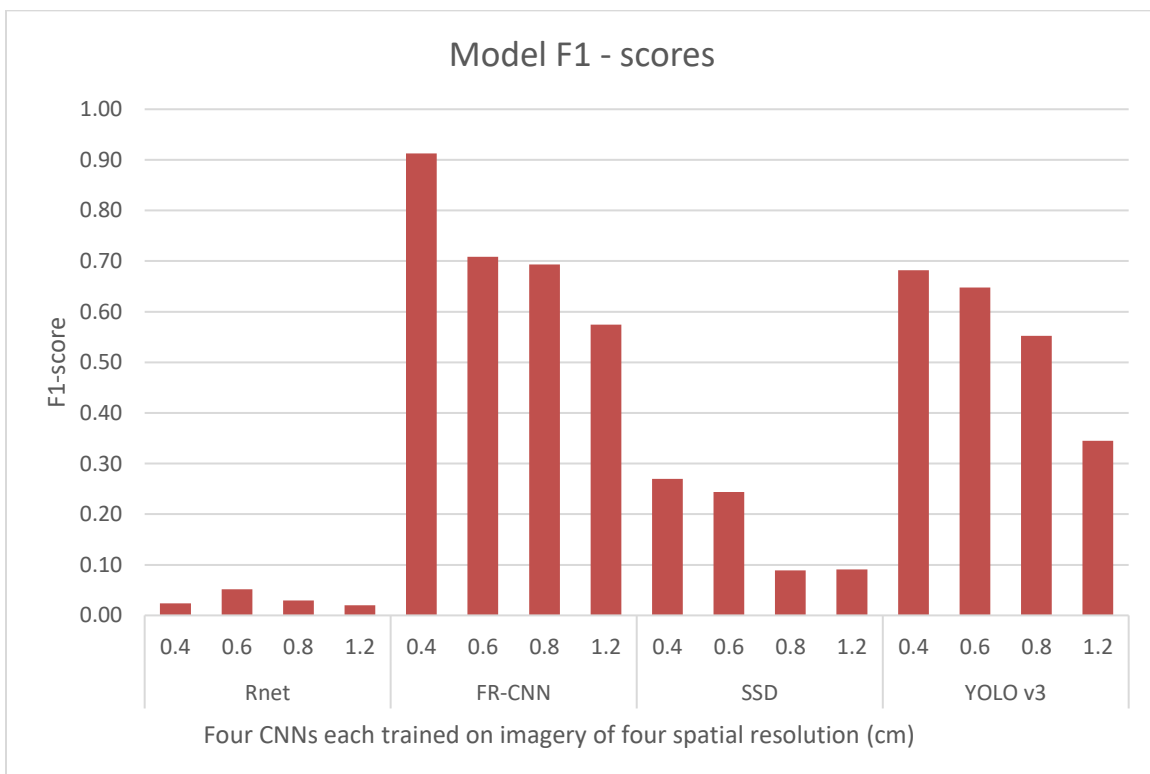|	RetinaNet	Faster R-CNN	SSD	YOLO v3
0.4 cm	0.02	0.91	0.27	0.68
0.6 cm	0.05	0.71	0.24	0.65
0.8 cm	0.03	0.69	0.09	0.55
1.2 cm	0.02	0.57	0.09	0.34

Figure 3.23: F1-scores results of 16 models evaluated on imagery from four production fields with 7,800 weeds.

Detection Rate

Detection times (min) and trial size (m^2) were used to calculate detection rates (ha/h) for each model in each of the four field deployments. Detection times were then averaged by model and the resulting means and standard deviations are summarized in Table 3.16 and Figure 3.23.

78

Null hypothesis four stated none of 16 custom trained open-source models (four CNNs trained on imagery of four spatial resolutions) will achieve mean detection rates (ha/h) of 4 ha/h. Seven of the 16 models produced a detection rate mean that exceeded 4 ha/h; therefore, null hypothesis four was rejected.

Mean detection rate of individual models consistently increased as training imagery spatial resolution increased. Comparing CNNs trained on the same spatial resolution imagery, RetinaNet produced the slowest detection rates except for at 0.8 cm GSD. Again, looking at CNNs trained on the same spatial resolution imagery, SSD consistently produced the highest detection rates. Faster R-CNN which produced the highest and most robust detection metrics (precision, recall, and F1-score) made these detections slower relative to YOLO v3 and SSD.

Table 3.16: Means and standard deviations of detection rates (ha/h) between CNNs and GSDs.

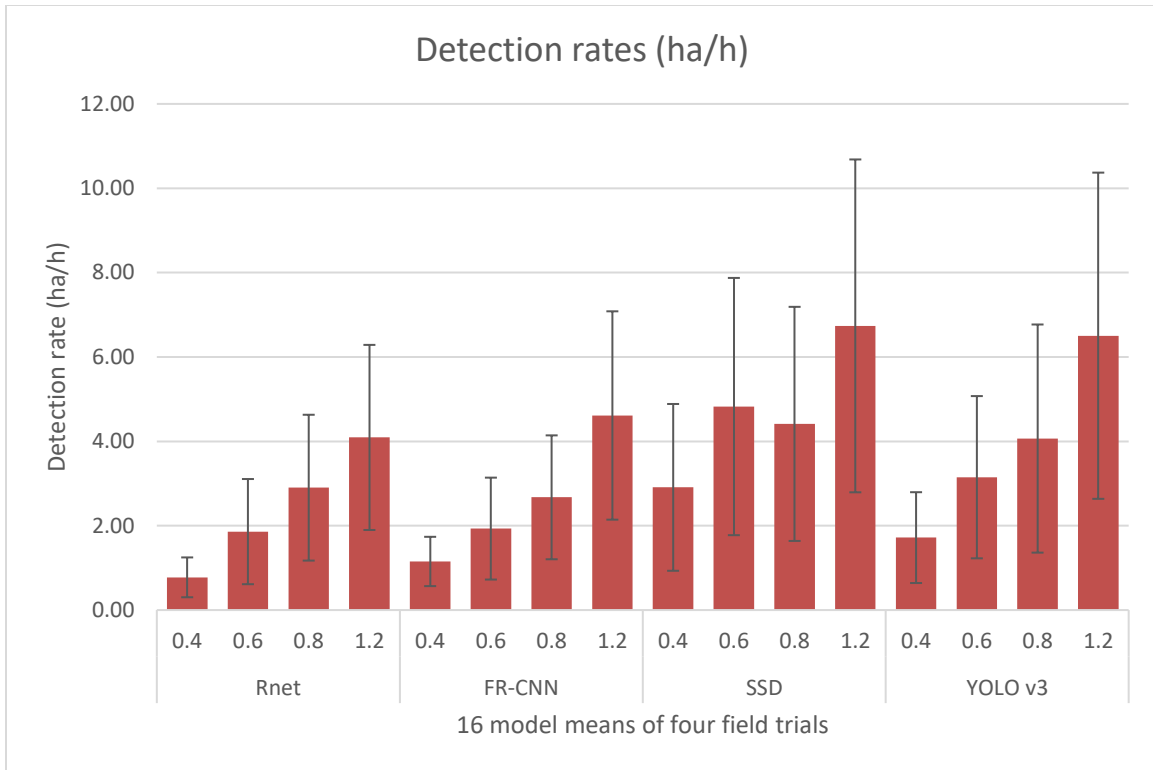|	RetinaNet		Faster R-CNN		SSD		YOLO v3	
	M	SD	M	SD	M	SD	M	SD
	--- (ha/h) ---							
0.4 cm	0.78	0.47	1.15	0.58	2.91	1.98	1.72	1.08
0.6 cm	1.86	1.25	1.93	1.21	4.82	3.05	3.15	1.92
0.8 cm	2.90	1.73	2.67	1.47	4.41	2.77	4.07	2.70
1.2 cm	4.09	2.19	4.61	2.47	6.74	3.95	6.50	3.87

Figure 3.24: Means and standard deviations of model detection rates (ha/h) of four field trials.

3.5   Discussion

The following section describes the results stated in chapter 3.4 and how they relate to

current production environments. Also discussed are the limitations of these results and practical

deployment of these algorithms in production systems.

3.5.1   Model precision

Null hypothesis one stated that none of 16 custom trained open-source models (four

CNNs trained on four spatial resolutions) will achieve $\geq$ 0.80 precision results at *Palmer*

amaranth detection in soybean field imagery. Initial evaluations using the subset of training

imagery withheld for testing indicated that three of the custom trained models achieved $\geq$ 0.80

precision scores including YOLO v3 trained on 0.4 cm spatial resolution imagery (precision =

0.92), YOLO v3 trained on 0.6 cm spatial resolution imagery (precision = 0.86), and Faster R-

CNN trained on 0.4 cm spatial resolution imagery (precision = 0.80). Additional deployments to four field trials evaluated model precisions on nearly 7,800 weeds. Only Faster R-CNN trained on 0.4 cm GSD maintained a precision value of ≥ 0.80 with a reported precision of 0.86. Therefore, null hypotheses one was rejected.

It was also noted that the CNN type heavily influenced precision results. RetinaNet produced an excessive number of false positives holding all its precision values across training imagery resolutions below 0.03. It is suspected that RetinaNet's designed ability to separate foreground and background complicated detections in imagery where equal numbers of potential detections were at both scales. Faster R-CNN performed the best when comparing CNNs across equivalent training imagery spatial resolutions (0.86, 0.59, 0.58, and 0.47 at 0.4, 0.6, 0.8, and 1.2 cm spatial resolutions, respectively). It was observed that in deployment field two that crop residues triggered false positives, reducing each model's relative precision. Likewise, relative precision results were also reduced in deployment field four due to false detections made on individual soybean plants visible from non-continuous rows. Each of these field examples represent opportunities for improvement in the training imagery for future models and limitations for field deployments in the current version models.

In the context of field performance when applied to selective herbicides, model precision is of secondary importance. Detection recall must be maintained as high as possible first and reductions through improvements of precision are considered. This consideration is different in the case of non-selective herbicides where damage can be caused by spraying desirable vegetation.

3.5.2 Model recall

Null hypothesis two stated that none of 16 custom trained open-source models (four

CNNs trained on imagery of four spatial resolutions) will achieve $\geq$ 0.95 recall results at *Palmer*

amaranth detection in soybean field imagery. Initial evaluations using the subset of training

imagery withheld for testing indicated that Faster R-CNN trained on 0.4 cm GSD imagery alone

exceeded this threshold and achieved a 0.96 recall score. Additional deployments four field trials

evaluated each model's recall on nearly 7,800 weeds. Again, only Faster R-CNN trained on 0.4

cm spatial resolution broke the 0.95 threshold and produced a recall value of 0.98. Therefore,

null hypotheses two was rejected.

CNN type heavily influenced recall results. SSD produced the lowest recall values never

exceeding a recall of 0.25 across all training imagery spatial resolutions. Faster R-CNN

performed the best when comparing CNNs across equal training imagery resolutions (0.96, 0.89,

0.87, and 0.74 at 0.4, 0.6, 0.8, and 1.2 cm spatial resolution imagery, respectively).

In the context of field performance when applied to selective herbicides, model recall is

of upmost importance and must be maintained as high as possible in a site-specific workflow.

Field level recall of 0.95 and greater as demonstrated by Faster R-CNN trained on 0.4 spatial

resolution is considered field ready and deployable in the context of these reported results.

Special consideration was given to recall and how it was evaluated with orthomosaic

derived keys. Deployment field four was used to evaluated orthomosaic derived key which

produced a recall of 0.83 for all weeds observed via *in situ* scouting; however, recall increased to

0.92, 0.96, and 0.97 as the weed size threshold was increased to $\geq$ 2.5, $\geq$ 5.0, and $\geq$ 7.5 cm,

respectively. A limiting threshold was observed in keying weeds in orthomosaic when they are

less than 2.5 cm in height. Larger weeds were easier to see, and subsequently received a higher recall score. This limitation would suggest that targeting weeds > 2.5 cm would obtain the highest detection efficacy when using 0.4 cm spatial resolution. It was observed by *in situ* scouting that field deployments one, two, and three contained larger weeds that escaped a previous herbicide application, while any smaller emerging weeds were suppressed by a preemergent herbicide applied at the same time.

3.5.3   F1-score

Null hypothesis three stated that none of 16 custom trained open-source models (four CNNs trained on imagery of four spatial resolutions) will achieve $\geq$ 0.90 F1-score at *Palmer amaranth* detection in soybean field imagery. Initial evaluations using the subset of training imagery withheld for testing showed that the highest F1-score achieved was by Faster R-CNN trained on 0.4 cm spatial resolution imagery (F1 = 0.87). Additional deployments in four field trials evaluated each model's precision and recall on nearly 7,800 weeds. Only Faster R-CNN trained on 0.4 cm spatial resolution imagery broke the 0.90 F1-score threshold producing a score of 0.91. Therefore, null hypotheses three was rejected.

Like precision and recall, CNN type heavily influenced F1-score results. RetinaNet produced the lowest F1-score values never exceeding F1-score = 0.05 across all GSDs. Faster R-CNN performed the best when comparing CNNs across equal GSDs (0.91, 0.71, 0.69, and 0.57 at 0.4, 0.6, 0.8, and 1.2 cm GSD, respectively). Except for at 0.4 cm GSD, YOLO v3 F1- scores closely followed Faster R-CNN at 0.68, 0.65, 0.55, and 0.34 (0.4, 0.6, 0.8, and 1.2 cm GSD, respectively).

As a harmonized mean of precision and recall, F1-scores reward or punish extreme values. In a site-specific selective herbicide application, a high F1-score achieved by a high recall value supplemented by a reasonably high precision value would be most advantageous. Field level recall of ≥ 0.95 in combination with an F1-score of ≥ 0.90 is considered field ready and deployable.

3.5.4   Detection rate

Null hypothesis four stated that none of 16 custom trained open-source models (four CNNs trained on imagery of four spatial resolutions) will achieve detection rates of 4 ha/h. This detection rate threshold was established with two figures in mind. In the U.S, it is common for fields to be in 16 or 32 ha units. If imagery were collected during a field scouting event, allowing four hours for detections to be completed on 16 ha or eight hours on 32 ha would allow for the overnight processing of imagery, producing results that can be applied the next day. 32 ha in eight hours or 16 ha in four hours is a detection rate of four ha/h. While model accuracy, properly balanced between recall and precision is essential, the detection rate of a model will influence a workflows scalability. Excessive detection rates for field scale applications will inhibit site-specific map generation and limit timely action on the results produced.

Results showed that the detection rates (ha/h) consistently decreased as spatial resolution of imagery increased. This was expected since higher spatial resolution imagery contains more pixels to convolve across each layer of the CNN. SSD trained on 1.2 cm spatial resolution imagery produced the highest mean detection rate when averaged across four field trials at 6.74 ha/h ($SD = 3.95$), closely followed by YOLO v3 trained on 1.2 cm spatial resolution imagery with a mean detection rate of 6.5 ha/h ($SD = 3.87$). The highest performing models based on model precision, recall, and F1-scores produced some of the slowest detection rates (Faster R-:

M = 1.15, *SD* = 0.58 and YOLO v3: *M* = 1.72, *SD* = 1.08, both trained on 0.4 cm spatial

resolution imagery).

It should be noted that these detection rates were produced on a laptop that would be

considered accessible to producers looking to deploy such a workflow; however, significantly

faster detection rates would be expected from hardware leveraging more powerful GPU

computing power optimized for object detection tasks.

3.6　　Conclusion

Current *in situ* scouting of crop production fields for the presence of weeds is laborious

and does not capture the in-field variability that is often present. Furthermore, the near future

availability of site-specific technologies for weed control will require the ability to accurately

differentiate between crops and weeds. Deep learning in the form of CNNs has the potential to

become a prominent analysis tool in agricultural image processing applications for their abilities

at finding objects such as weeds in imagery that would not be practical for a human observer to

detect at field scales.

Sixteen CNN models (RetinaNet, Faster R-CNN, YOLO v3, and SSD each trained on

0.4, 0.6, 0.8, and 1.2 cm spatial resolution imagery) were trained to detect *Palmer amaranth* in

soybean field imagery. Field deployments of models were conducted and model precision, recall,

F1-scores, and detection rates were measured in imagery containing 7,800 weeds from four

production fields of slightly different crop maturity, weed size and density, plant and soil color.

Results showed that CNN architecture and training imagery resolution influenced model results.

Detection accuracies increased proportionally, and detection rates decreased proportionally with

increasing spatial resolution of training imagery. Of the 16 models trained, only the combination

of Faster R-CNN trained on 0.4 cm spatial resolution imagery proved to meet field ready expectations for precision, recall and F1-score. While Faster R-CNN maintained impressive detection metrics (80% precision, 98% recall, 94% F1-scores) across deployments in four trials over two years, further testing of its ability to generalize is recommended. Specifically, an increase in false positive detections were noted at the edges of fields where soybeans were stunted and smaller, when soybean rows were broken instead of continuous, and when excessive crop residue was present on the soil. This reduced the precision of some fields and suggested that accuracies in these field conditions may need to be assessed further or adjusted in future training efforts. Furthermore, detection recall of weed heights smaller than 2.5 cm fell significantly compared with weed heights $\geq$ 2.5 cm. This limitation suggests that targeting weeds that escape control of a previous herbicide application may be an optimal use case for any site-specific workflow leveraging 0.4 cm spatial resolution imagery. Detection rates were determined to be marginally practical on coarser imagery; however, high resolution imagery detection times became excessive in the context of a single day turnaround timeline with the computing hardware used in this experiment.

3.7   References

Boulent, J., Foucher, S., Théau, J., & St-Charles, P.-L. (2019). Convolutional Neural Networks for the Automatic Identification of Plant Diseases. *Frontiers in Plant Science*, *10*, 941. https://doi.org/10.3389/fpls.2019.00941

Dobbels, A. A., & Lorenz, A. J. (2019). Soybean iron deficiency chlorosis high-throughput phenotyping using an unmanned aircraft system. *Plant Methods*, *15*(1), 97. https://doi.org/10.1186/s13007-019-0478-9

Fernández-Quintanilla, C., Peña, J. M., Andújar, D., Dorado, J., Ribeiro, A., & López-Granados, F. (2018). Is the current state of the art of weed monitoring suitable for site-specific weed management in arable crops? *Weed Research*, *58*(4), 259–272. https://doi.org/10.1111/wre.12307

Gai, J., Tang, L., & Steward, B. L. (2020). Automated crop plant detection based on the fusion of color and depth images for robotic weed control. *Journal of Field Robotics*, *37*(1), 35–52. https://doi.org/10.1002/rob.21897

Gaston, L. A., Locke, M. A., Zablotowicz, R. M., & Reddy, K. N. (2001). Spatial Variability of Soil Properties and Weed Populations in the Mississippi Delta. *Soil Science Society of America Journal*, *65*(2), 449–459. https://doi.org/10.2136/sssaj2001.652449x

Gianessi, L. P., & Reigner, N. P. (2007). The Value of Herbicides in U.S. Crop Production. *Weed Technology*, *21*(2), 559–566. https://doi.org/10.1614/WT-06-130.1

He, X., Cui, T., Zhang, D., Wei, J., Wang, M., Yu, Y., Liu, Q., Yan, B., Zhao, D., & Yang, L. (2017). Development of an electric-driven control system for a precision planter based on a closed-loop PID algorithm. *Computers and Electronics in Agriculture*, *136*, 184–192. https://doi.org/10.1016/j.compag.2017.01.028

Khan, S., Tufail, M., Khan, M. T., Khan, Z. A., & Anwar, S. (2021). Deep learning-based identification system of weeds and crops in strawberry and pea fields for a precision agriculture sprayer. *Precision Agriculture*, *22*(6), 1711–1727. https://doi.org/10.1007/s11119-021-09808-9

Lambert, J., Childs, D., & Freckleton, R. (2019). Testing the ability of unmanned aerial systems and machine learning to map weeds at subfield scales: A test with the weed *ALOPECURUS MYOSUROIDES* (Huds). *Pest Management Science*, ps.5444. https://doi.org/10.1002/ps.5444

Lambert, J., Hicks, H. L., Childs, D. Z., & Freckleton, R. P. (2018). Evaluating the potential of Unmanned Aerial Systems for mapping weeds at field scales: A case study with *Alopecurus myosuroides*. *Weed Research*, *58*(1), 35–45. https://doi.org/10.1111/wre.12275

Li, M., Zhang, Z., Lei, L., Wang, X., & Guo, X. (2020). Agricultural Greenhouses Detection in High-Resolution Satellite Images Based on Convolutional Neural Networks: Comparison of Faster R-CNN, YOLO v3 and SSD. *Sensors*, *20*(17), 4938. https://doi.org/10.3390/s20174938

Liang, Q., Xiang, S., Hu, Y., Coppola, G., Zhang, D., & Sun, W. (2019). PD2SE-Net: Computer-assisted plant disease diagnosis and severity estimation network. *Computers and Electronics in Agriculture*, *157*, 518–529. https://doi.org/10.1016/j.compag.2019.01.034

Lin, Wu, Fu, Wang, Zhang, & Kong. (2019). Dual-NMS: A Method for Autonomously Removing False Detection Boxes from Aerial Image Object Detection Results. *Sensors*, *19*(21), 4691. https://doi.org/10.3390/s19214691

López-Granados, F. (2011). Weed detection for site-specific weed management: Mapping and real-time approaches: Weed detection for site-specific weed management. *Weed Research*, *51*(1), 1–11. https://doi.org/10.1111/j.1365-3180.2010.00829.x

Lottes, P., Behley, J., Chebrolu, N., Milioto, A., & Stachniss, C. (2020). Robust joint stem detection and crop-weed classification using image sequences for plant-specific treatment in precision farming. *Journal of Field Robotics*, *37*(1), 20–34. https://doi.org/10.1002/rob.21901

Lottes, P., Hörferlin, M., Sander, S., & Stachniss, C. (2017). Effective Vision-based Classification for Separating Sugar Beets and Weeds for Precision Farming: Effective Vision-Based Classification. *Journal of Field Robotics*, *34*(6), 1160–1178. https://doi.org/10.1002/rob.21675

Lou, X., Huang, Y., Fang, L., Huang, S., Gao, H., Yang, L., Weng, Y., & Hung, I.-K. uai. (2022). Measuring loblolly pine crowns with drone imagery through deep learning. *Journal of Forestry Research*, *33*(1), 227–238. https://doi.org/10.1007/s11676-021-01328-6

Louargant, M., Villette, S., Jones, G., Vigneau, N., Paoli, J. N., & Gée, C. (2017). Weed detection by UAV: Simulation of the impact of spectral mixing in multispectral images. *Precision Agriculture*, *18*(6), 932–951. https://doi.org/10.1007/s11119-017-9528-3

Mahlein, A.-K. (2016). Plant Disease Detection by Imaging Sensors – Parallels and Specific Demands for Precision Agriculture and Plant Phenotyping. *Plant Disease*, *100*(2), 241–251. https://doi.org/10.1094/PDIS-03-15-0340-FE

Milioto, A., Lottes, P., & Stachniss, C. (2018). Real-time Semantic Segmentation of Crop and Weed for Precision Agriculture Robots Leveraging Background Knowledge in CNNs. *ArXiv:1709.06764 [Cs]*. http://arxiv.org/abs/1709.06764

Nguyen, K., Huynh, N. T., Nguyen, P. C., Nguyen, K.-D., Vo, N. D., & Nguyen, T. V. (2020). Detecting Objects from Space: An Evaluation of Deep-Learning Modern Approaches. *Electronics*, *9*(4), 583. https://doi.org/10.3390/electronics9040583

Pantazi, X. E., Tamouridou, A. A., Alexandridis, T. K., Lagopodi, A. L., Kashefi, J., & Moshou, D. (2017). Evaluation of hierarchical self-organising maps for weed mapping using UAS multispectral imagery. *Computers and Electronics in Agriculture*, *139*, 224–230. https://doi.org/10.1016/j.compag.2017.05.026

Sishodia, R. P., Ray, R. L., & Singh, S. K. (2020). Applications of Remote Sensing in Precision Agriculture: A Review. *Remote Sensing*, *12*(19), 3136. https://doi.org/10.3390/rs12193136

Tamouridou, A., Alexandridis, T., Pantazi, X., Lagopodi, A., Kashefi, J., Kasampalis, D., Kontouris, G., & Moshou, D. (2017). Application of Multilayer Perceptron with Automatic Relevance Determination on Weed Mapping Using UAV Multispectral Imagery. *Sensors*, *17*(10), 2307. https://doi.org/10.3390/s17102307

Tan, L., Huangfu, T., Wu, L., & Chen, W. (2021). Comparison of RetinaNet, SSD, and YOLO v3 for real-time pill identification. *BMC Medical Informatics and Decision Making*, *21*(1), 324. https://doi.org/10.1186/s12911-021-01691-8

Toda, Y., & Okura, F. (2019). How Convolutional Neural Networks Diagnose Plant Disease. *Plant Phenomics*, *2019*, 1–14. https://doi.org/10.34133/2019/9237136

Van Wychen, L. (2016). *2015 Baseline survey of the most common and troublesome weeds in the United States and Canada*. https://wssa.net/wp-content/uploads/WSSA-Weed-Survey-2017-05-19.pdf

Van Wychen, L. (2017). *2016 WSSA Survey Ranks Most Common and Most Troublesome Weeds in Broadleaf Crops, Fruits and Vegetables*. https://wssa.net/wp-content/uploads/WSSA-Weed-Survey-2017-05-19.pdf

Van Wychen, L. (2021). *2021 WSSA Survey of the Most Common and Troublesome Weeds in Aquatic and Non-Crop Areas in the United States and Canada.* Weed Science Society of America National Weed Survey Dataset. https://wssa.net/wp-content/uploads/2021-survey-aquatic-noncrop.xlsx

Varela, S., Dhodda, P., Hsu, W., Prasad, P. V., Assefa, Y., Peralta, N., Griffin, T., Sharda, A., Ferguson, A., & Ciampitti, I. (2018). Early-Season Stand Count Determination in Corn via Integration of Imagery from Unmanned Aerial Systems (UAS) and Supervised Learning Techniques. *Remote Sensing*, *10*(3), 343. https://doi.org/10.3390/rs10020343

Vasconez, J. P., Delpiano, J., Vougioukas, S., & Auat Cheein, F. (2020). Comparison of convolutional neural networks in fruit detection and counting: A comprehensive evaluation. *Computers and Electronics in Agriculture*, *173*, 105348. https://doi.org/10.1016/j.compag.2020.105348

Wang, A., Zhang, W., & Wei, X. (2019). A review on weed detection using ground-based machine vision and image processing techniques. *Computers and Electronics in Agriculture*, *158*, 226–240. https://doi.org/10.1016/j.compag.2019.02.005

4.0   SITE-SPECIFIC WEED APPLICATIONS USING PWM SPRAYERS

Abstract

Weed populations in agricultural production fields are often scattered in their distribution and few in numbers compared with the planted crop; however, herbicides are broadcast across fields evenly. Although effective, in the case of post-emergent herbicides, this method often uses exceedingly more pesticides than necessary. A novel weed detection and control workflow was evaluated targeting *Palmer amaranth* in soybean (*Glycine max)* fields*.* A site-specific workflow leveraging UAS imagery, custom trained CNN models for weed detection, and applied by a pulse width modulated (PWM) sprayer was evaluated in replicated field trials established in 2021 and 2022. Weed control (%) was compared between a broadcast treatment and the proposed site-specific workflow. Results indicate no statistical ($p < .05$) difference in weed control measured one ($M = 96.22\%$, $SD = 3.90$ and $M = 90.10\%$, $SD = 9.96$), two ($M = 95.15\%$, $SD = 5.34$ and $M = 89.64\%$, $SD = 8.58$), and three weeks (M = 88.55, SD = 11.07 and M = 81.78%, SD = 13.05) after application between broadcast and site-specific treatments, respectively. Furthermore, there was a significant ($p < 0.05$) 48% mean reduction in applied area (m^2) between broadcast and site-specific treatments across both years. Equivalent application efficacy can be achieved with significant reductions in herbicides if weeds are targeted through site-specific applications. Site-specific weed maps can be generated and executed using accessible technologies like UAS, open-source CNNs, and PWM sprayers.

4.1   Introduction

Continuous advances in sensor, computing, and positioning technologies have encouraged exploration of automated systems in agriculture (Gonzalez-de-Soto et al., 2016).

Aravind et al. (2017) generally grouped these novel "smart" farming systems into five categories including tillage, soil/nutrient analysis, seeding/transplanting, scouting/control measures, and harvesting. Applications in most of these categories have been demonstrated in both small novel advanced autonomous robots and in modifications to current field equipment. Each of these novel robotic systems represent incremental improvements and advancements toward truly autonomous agricultural systems. However, more field ready applications within these categories generally leverage the power and robustness of manned equipment designed for field conditions with modifications to capabilities or actions informed by an automated system.

Soil tillage is underrepresented in small automated robotic research due to the heavy lifting nature of the practice that is more appropriately handled by large equipment (Aravind et al., 2017); however, large equipment advancements are represented. Akhalaya & Shogenov (2018) demonstrated an automated multistage tillage implement where the depth obtained by a previous cultivation stage, informs, and adjusts the depth of subsequent stages accordingly in real-time. The field ready and robust design was mounted behind a tractor and consisted of a series of magnetic limit switches informing hydraulic systems adjustments. The authors claimed a 10% increase in tillage quality compared to a "one depth" fits all. While this implements "information system" may be relatively simple, it is this simplicity that allows real-time information to be tuned into real-time site-specific action independent of operator input.

Soil nutrient sampling has been attempted with small robotics. Kitić et al. (2022) developed and evaluated an autonomous robotic soil sampler and analyzer – Agrobot Lala. The robot collects, georeferences, and analyzes *in situ* soil samples. It then generates variable rate fertilizer maps that are uploaded to a cloud-based platform for application via manned equipment. The authors noted a successful deployment of the robot at the plot level with further testing need for

scalability to field level applications. Much work has also been completed using remote sensing techniques to map in field nutrient needs, with nitrogen management of particular importance (Corti et al., 2019). Corti et al. (2019) mapped the infield nitrogen variability using a high resolution green adjusted normalized difference vegetative index (GNDVI) collected via UAS imagery. Nitrogen content was successfully estimated with a regression model reporting a $R^2 =$ .89. Basso et al. (2016) demonstrated that site-specific nitrogen application maps could be generated using vegetative indices calculated from high resolution satellite imagery. Field testing showed that remotely sensed imagery highly correlated with *in situ* soil measurements. Additionally, the team applied the site-specific maps using large field equipment and determined that nitrogen use efficiency was shown to increased when nitrogen was varied according to measured needs.

Automated and site-specific planting efforts have been explored. Tatsuno et al. (2022) demonstrated a transplanting robot that reported 84% success at transplanting of chain pot seedlings in shaft tillage cultivation. While the transplantation step was completed automatically, the driving and navigation of the system was done manually by the researchers pushing a cart through the plots. Field ready equipment have been used to demonstrate site-specific seeding with consumption potatoes where seed spacing was varied based on fertility indicators from remote sensing and *in site* measurements. Site-specific seeding informed by high-resolution fertility maps was compared with traditional broadcast seeding and results indicate improved tuber yields, savings on seeds, and higher net profits for site-specific treatments compared to traditional broadcasting methods.

Harvesting of high values crops has been demonstrated with varied success by small, automated robots. Lehnert et al. (2020) claimed "performance improvement" in automated

harvesting of sweet pepper using a novel robot with CNN detection capabilities, producing a

76.5% success rate. Similarly, Oktarina et al. (2020) developed a robot vision systems and arm

capable of detecting and harvesting green and red tomatoes and tested in laboratory conditions.

The team did not report the detection accuracy of the vision system, instead emphasizing the that

the full cycle of steps from detection, harvesting, placement of fruit, and ready to detect again

averaged 9.68 and 10.59 seconds for green and red tomatoes respectively. Field scale testing of

harvesting equipment has also been demonstrated. Optimized harvesting routes were generated

to inform a commercially available autonomous field ready robot (Robotti) harvester, based on

higher protein values in wheat fields (Villa-Henriksen et al., 2021). High protein value areas

were harvested separately (site-specific) than lower protein areas allowing for higher value

marketing for portions of the crop. Results suggest that route and harvest optimization was

effective; however, the economic returns were often negated from extra time in the field

harvesting. Roshanianfard & Noguchi (2020) designed and tested a unique pumpkin harvesting

"end-effector" implement. The specialized arm and vision system were mounted on the front of a

manned field tractor. Evaluation of the systems suggested it had a harvest possibility of 92% and

success rate of 79%. The team emphasized the designs versatility at generalizing to various size

and positioned pumpkins.

Scouting and control efforts have been demonstrated in weed detection and applied

mechanical or chemical controls applied. Gai et al. (2020) demonstrated the in-field real-time

classification of weeds among broccoli and lettuce crops. The team reported an overall 91%

accuracy at weed detection. Ruigrok et al. (2020) evaluated a site-specific pesticide application

workflow in potato fields using an autonomous field robot reporting 96% effective control over

weeds. Utstumo et al. (2018) deployed an autonomous field sprayer to site-specifically treat

weeds in a carrot crop reporting a ten-fold reduction in herbicide use with 100% weed control.

While robotic mechanical weeders and sprayers are an impressive fully autonomous goal,

like other agricultural practices, detection and treatment are not easily scalable with small

robotics. In the case of pesticide applications, treatment rates are as high as 320 m^2/sec. Great

improvements in system efficiencies are necessary before this technology can be used for

selective herbicide applications on-site and in real-time (Fernández-Quintanilla et al., 2018).

Similar to the other four categories that Aravind et al. (2017) outlined, the semi-

automation of scouting and control of pests at the field level is likely closer to field readiness

when currently adopted equipment is deployed and either modified or informed by site-specific

maps. As pesticide application equipment has evolved to become more accurate at holding rates

constant, some efforts have been made to leverage these capabilities to selectively treat areas.

Castaldi (2017) used imagery collected from a UAS to develop prescription maps and spot treat

patches of weeds in corn fields using field equipment with section control (two-meter sections).

Their results showed nearly a 40% decrease in herbicide used when selectively spraying with no

significant difference in the crop or weed biomass compared to the broadcast treatment.

Similarly, Campos (2020) used MS imagery captured from a UAS to develop prescription spray

maps for vineyards. Imagery in combination with sample *in situ* canopy measurements mapped

canopy density allowing pesticide rate to be varied across the field as site-specific needs varied.

Results showed similar control of the target pest with reduced rates of water volume and total

pesticides applied. Both Castaldi et al. (2017) and Campos et al. (2020) leveraged field

equipment with section control common to modern production sprayers which use electronically

actuated shut-off valves to manage turning on and shutting off groups of nozzles often meters

upstream from the nozzle in plumbing. The actuation speed and section resolutions of these systems are not ideal for site-specific application and marginally acceptable for treatment of sections or zones within fields (Luck et al., 2011).

PWM sprayers contain electronically actuated high-speed solenoids immediately upstream from the spray nozzle which control the flow rate and pressure of each nozzle independently. Research on PWM in field applications has shown that systems are highly responsive and can provide more flexible options for applying pesticides than traditional sprayers (Butts, et al., 2019). This research has primarily focused on the ability of PWM sprayers to maintain even pesticide coverage (rate and droplet size) among variable application speeds. While not designed for site-specific applications, the characteristic of PWM systems that enable stable applications amongst variable application conditions are the mechanical foundations for varying the treatment in a stable application environment. The potential use of PWM in site-specific applications represents improvement in the spatial resolution of treatment maps that limited both Castaldi et al. (2017) and Campos et al. (2020). Additionally the standard actuation control of 10 hz at the nozzles is a notable improvement from the research conducted by Luck et al. (2011). The evaluation of PWM sprayers deployed in site-specific herbicide applications has been largely unstudied. This work will use treatment maps generated from a UAS imagery and CNN detection workflow to inform a site-specific treatment applied by a PWM sprayer.

4.2   Study site

This research was conducted at the Jackson County Extension Center (35.5715763°N, 91.2539575°W) near Newport, Arkansas, USA in 2021 (year one) and 2022 (year two). Trials were located within subsections of two different production soybean fields, both planted in Enlist

E3™ soybeans (Corteva, 9330 Zionsville Road, Indianapolis, IN 46268) (Figure 4.1). Soybean

rows were planted at 76 cm spacing, Year one soybeans on treatment date were at V6 growth

stage (six trifoliate leaves) (planting date June 16th sprayed July 26th, 32 DAE, irrigated). Year

two soybeans on treatment date were at R1 (beginning flower) growth stage (planting date June

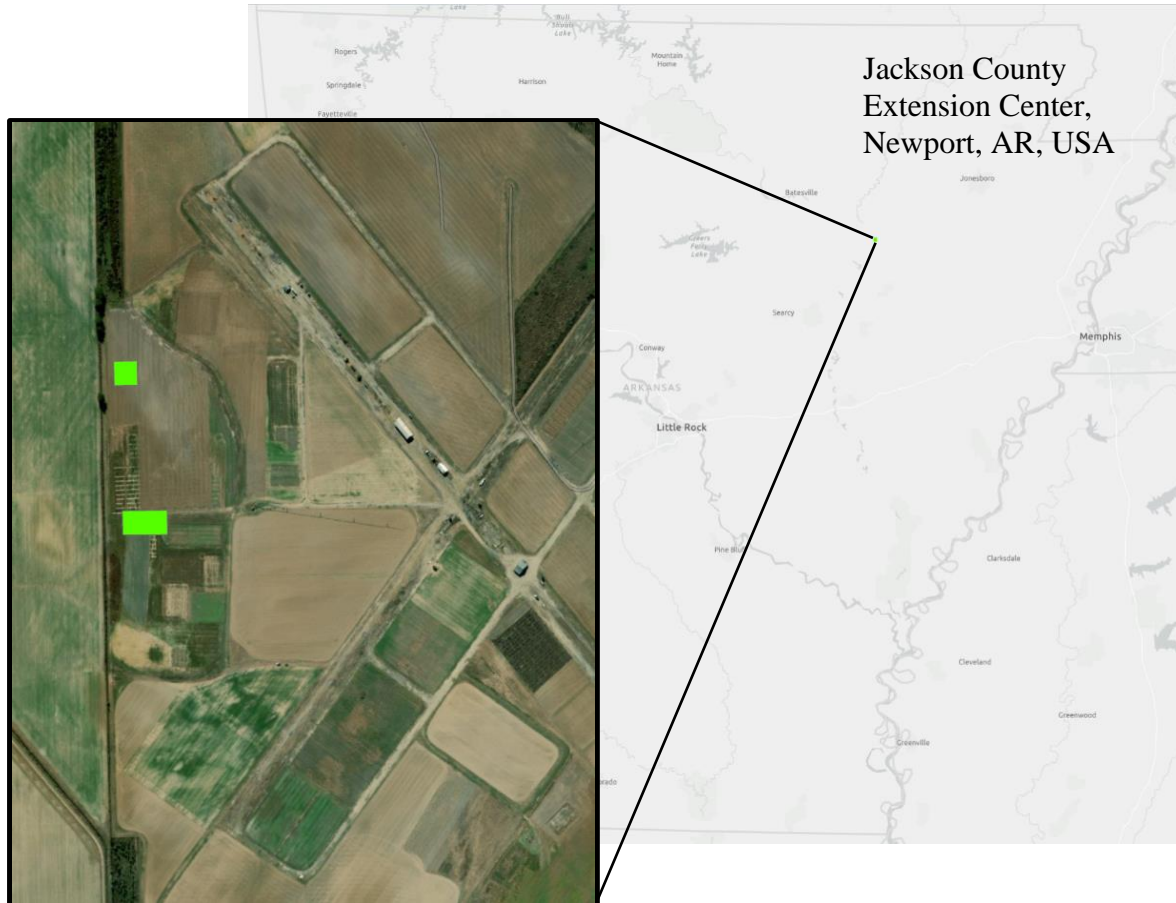9th, sprayed July 27th, 38 DAE, dryland).

Figure 4.1: Two study sites (2021 and 2022) at the Jackson County Extension Center, Newport, Arkansas, USA.

4.3   Methods

4.3.1   Experimental design

This study was an experimental design testing one independent variable- herbicide treatment method with three factors, untreated check (UTC), broadcast, and site-specific applications. Two randomized, replicated block experiments were conducted with plots three-meters-wide, 30-meter-long in year one and 20-meter-long in year two. Year one trial included six replications and year two trial included eight replications. The treatment map generation used inform site-specific applications are described in section 4.3.2. Both broadcast and herbicide treatments contained glufosinate-ammonium at 656 g ha−1 (Liberty® 280 SL, Bayer CropScience LP, Research Triangle Park, NC 27709).

One dependent variable was measured, individual weed control with ratings each taken seven-, 14-, and 21-day post treatment. Individual weed ratings within each plot were averaged to produce a per plot weed control mean for each collection date. RGB UAS imagery was captured at 0.4 cm GSD one day and five days prior to treatment, (year one and two respectively) and seven-, 14-, and 21-day post treatment in both years.

4.3.2   Site-specific treatment map creation

Real-time kinematic (RTK) global navigation satellite system (GNSS) ground control (GCPs) and control points (CPs) were established and maintained prior to each imagery data collection and in photogrammetric processing (Figure 4.2). GCPs were designated with marking paint and precise location (two cm) established using an RTK2B (Ardusimple Lleida, Spain) GNSS rover that leveraged a U-blox ZED-F9P microchip and a survey grade IP66 multiband antenna. Real-time kinematic (RTK) corrections were received through the Continuously

Operating Reference Station (CORS) network maintained by the Arkansas Department of Transportation (ARDOT) and broadcast through a Networked Transport of RTCM via Internet Protocol (NTRIP) caster. The CORS base station corrections used were from the ARDOT District Five headquarters outside of Batesville, Arkansas, USA approximately 39 km away from the study sites.
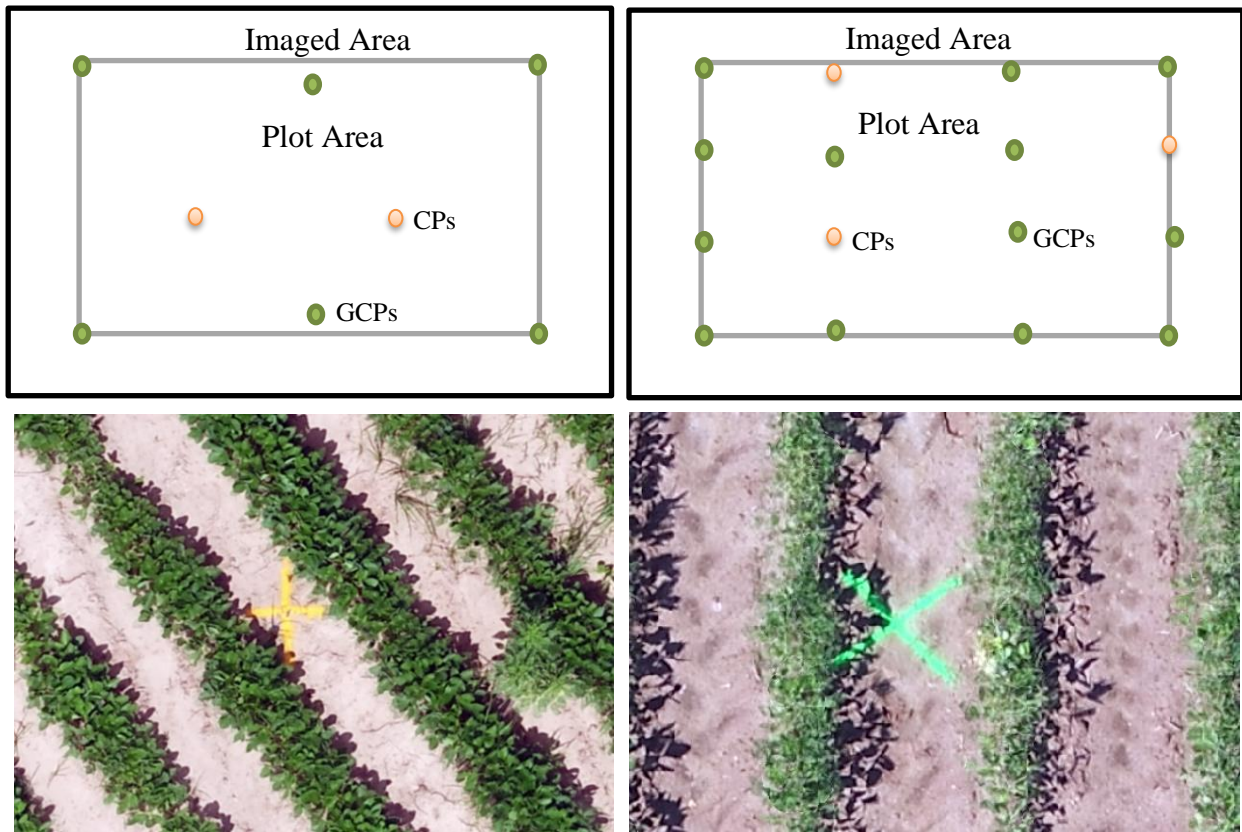

Figure 4.2: GCP layout in field trials for year one and two with marked GCP example.

Imagery for model training and field map generation was collected with a DJI Zenmuse X5s RGB sensor (iFlight Tech Company Ltd. Shenzhen, China) with the DJI micro 4/3 (MFT) lens (iFlight Tech Company Ltd. Shenzhen, China). This sensor and lens configuration has a 4/3" CMOS, 72° FOV, 15 mm focal length, f/1.7 focal ratio, and focus to infinite. Image

resolution of the camera is 5280 × 3956 (20.8 MP). The X5s sensor was mounted onto a DJI

Inspire II (iFlight Tech Company Ltd. Shenzhen, China) quadcopter (Figure 4.3).

Figure 4.3: DJI Inspire II and Zenmuse X5s (iFlight Tech Company Ltd. Shenzhen, China) UAS used to collect plot imagery.

All UAS flights were conducted in class G airspace and by a licensed part 107 pilot in

compliance with Federal Aviation Administration (FAA) regulations. For each training imagery

and field trial imagery collection, a unique flight plan was created, prior to take-of, within the

mobile application Pix4D Capture to facilitate image capture for processing in Pix4Dmapper.

Imagery was captured at 18.3 m altitude, 1.83 m/s flight velocity, 80% front overlap, and 70%

side overlap for each flight. At this altitude, the resulting imagery was at the targeted 0.4 cm

spatial resolution. As part of the Pix4D photogrammetric processing, the orthomosaic was

georeferenced and post corrections were performed using RTK collected GCPs.

During collection events, the UAS was manually controlled at take-off, ascended to target

altitude, and navigated into the target area using the mobile application DJI GO 4. While

hovering at altitude, the X5s's focus, and exposure was optimized for each flight using the

automated procedure within the DJI application where the operator taps on objects displayed on

the controller screen and the software adjusts sensor aperture and focus accordingly. The DJI

application was then closed, and the automated flight plan was initiated in the Pix4D Capture application.

Orthomosaic quality

Pix4Dmapper was used to co-register, georeference, and orthorectify the raw UAS image stream. The "Ag RGB" template was used for processing and its key default parameters are shown in Table 4.1. GCPs established in the field were used after initial processing and before point cloud generation to correct any geolocation errors. Rectified and georeferenced field orthomosaic (0.4 cm spatial resolution) was imported into ArcGIS Pro 2.8.2 (Esri inc. Redlands, California) for spatial analysis.

Table 4.1: Pix4Dmapper key parameters used to process image streams in "AG RGB" template.

Parameter	Value
1. Initial Processing	
Keypoint image scale	Full Scale Image
2. Point Cloud and Mesh	
Image Scale	¼, multiscale
Point Density	Low
Minimum number of matches	3
3. DSM, Orthomosaic, Index	
Resolution	1 x GSD
Orthomosaic	Merge GeoTIFF Tiles

Orthomosaic geolocation accuracy was assessed as part of orthomosaic processing. The quality reports provided by Pix4DMapper was used to compare both the accuracy of GCPs used in the stitching process as well as the CPs set aside for assessment. The X, Y, and Z axes root mean square (RMS) errors of for individual GCPs and CPs were assessed for each trial.

CNN models

Two custom trained CNNs were used to make weed detections on year one and year two trials. While this experiment was not intended to evaluate model performance specifically, the model performance does impact the overall workflow and ultimately the herbicide treatment efficacy. Therefore, the methods used to develop the custom models will be described and its detection accuracies reported in the following results section.

The two models used for field trials were YOLO v3 in year one and Faster R-CNN in year two, each trained using a custom image set containing two classes of 5,192 weeds (*Palmer amaranths* and broadleaf weeds) labeled in 0.4 cm spatial resolution imagery. ArcGIS Pro was used to train each of the models through the "Train Deep Learning Model (Image Analysis tools)". The experimental platform configuration used for training both models was the following: OS, Windows 10; GPU NVIDIA Quadro RTX 3000 with Max-Q Design; CPU, Intel® Core™ i7-1065 G7 CPU @ 1.30 GHz with 32 Gb of RAM. The backbone models Darknet-53 and Resnet-50 was used for YOLO v3 and Faster R-CNN, respectively. Both backbone models were trained on the ImageNet data set containing over one million images and were of similar size measured by their convolutional layers (50 and 53 layers for Resnet-50 and DarkNet-53, respectively). Learning rate was automatically determined for each model within the training tool and was 0.0002, 0.0020 and 0.00004, 0.00004 for YOLO v3 and Faster R-CNN, respectively. Batch size was limited by dedicated GPU memory on the training hardware and was held constant at two. Max epochs were set to 20. The detection precision, recall and F1-scores of trial specific model results were determined by calculating the differences between model outputs and annotated field imagery using an IoU of 0.1. These detection accuracies for trial specific models are reported in the following results section.

Site-specific map generation

ArcGIS Pro's ModelBuilder automated a series of steps generating a weed map layer (Figure 4.4). Field imagery processed into an orthomosaic raster and one of the custom trained deep learning models were used as input parameters. Additionally, an adjustable buffer distance around weeds (m), field application equipment resolution (m), and output folder are used to facilitate treatment map generation. The "Detect Objects with Deep learning" tool was used to process the detection threshold was set to 0.05.
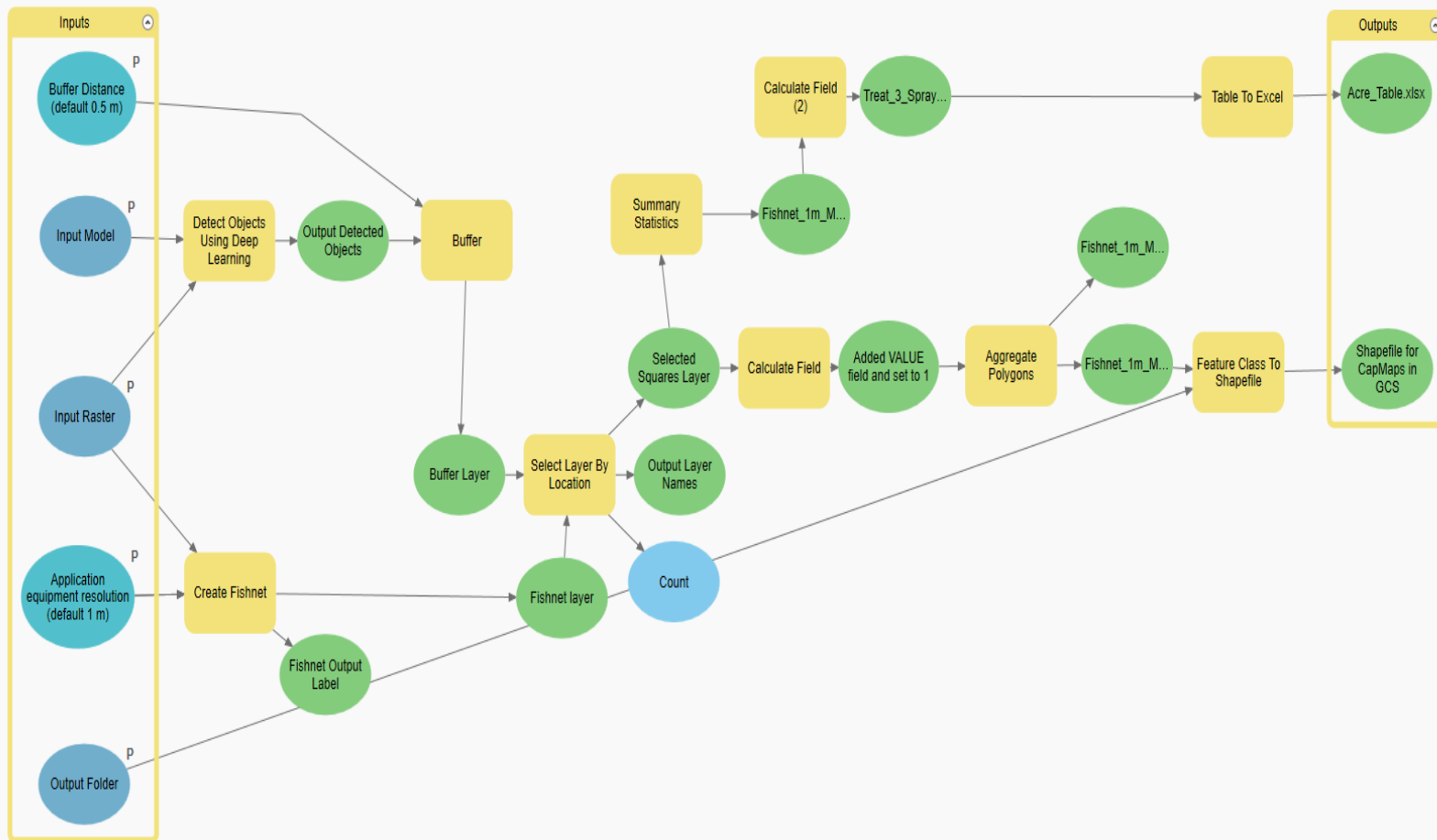
Figure 4.4: ModelBuilder used to generate weed map for field trials.
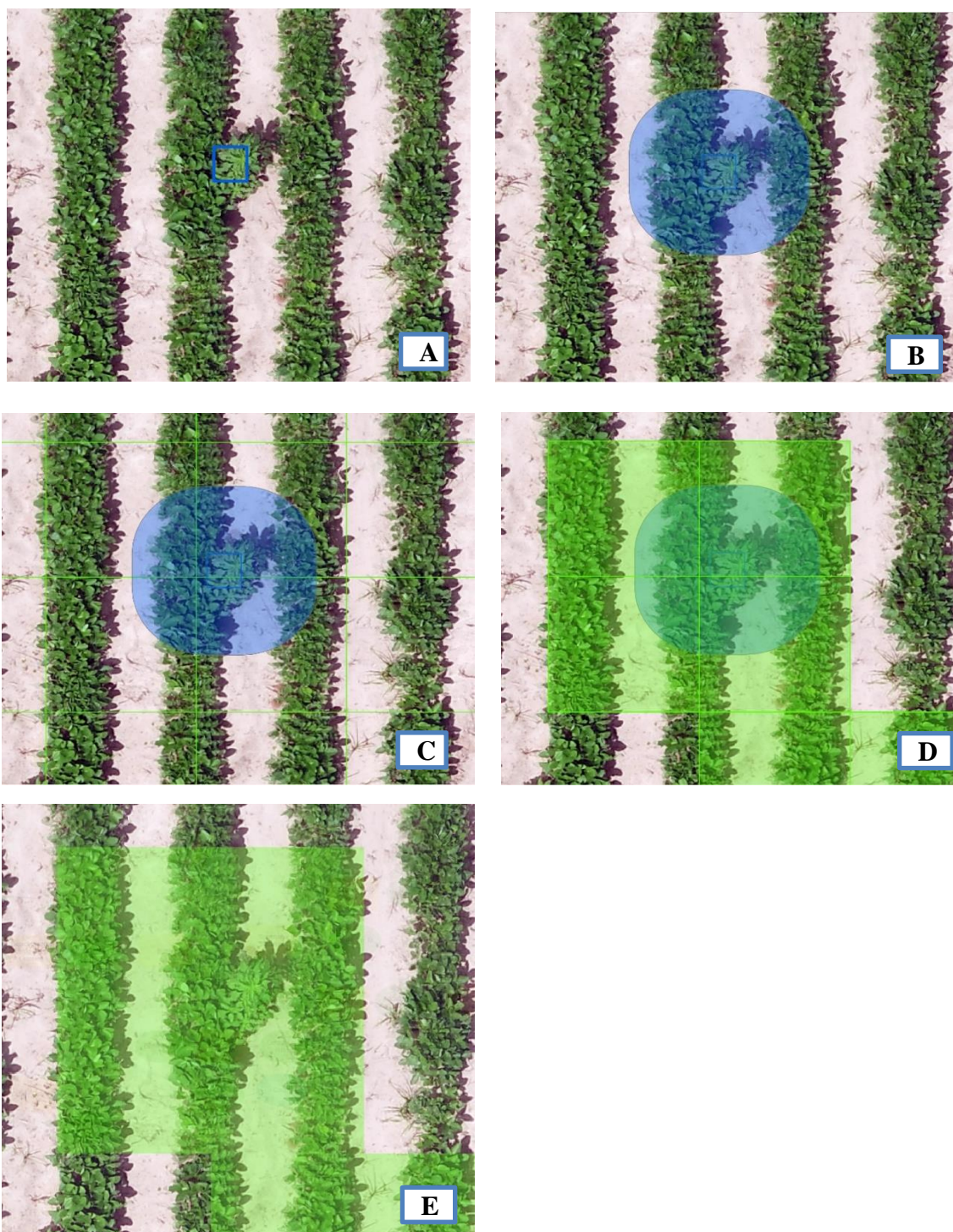
Figure 4.5: Example ModelBuilder steps to produce weed map. Detection box (A), applied 0.5 m buffer (B), applied 1 m fishnet (C), and selected 1 m grid boxes (D), and aggregate 1 m grid boxes (E).

A 0.5 m radius buffer was generated around each predicted bounding box and a one-meter fishnet grid squared and oriented north and south overlaid the extent of the input raster (plot area) (Figure 4.5). This combination ensured that weeds detected near or on the edge of a one-meter grid square would trigger the grid square adjacent as well. While this decreased the precision of the spray by unnecessarily treating more weed free areas, it also increased recall by ensuring that detected weeds were not missed because of their location relative to gridlines. The "Select by location" tool was used to select the grid squares that touched any weed buffers. Selected grid cells were then aggregated so that groups of selected cells were a single polygon. The selected and aggregated features were exported as a new layer and additionally their metadata was edited by adding a "value" column and populating all cells with "1". This acted as an ON signal in subsequent steps. Finally, the projected coordinates system "WGS_1984_UTM_Zone_15N" was transformed into the geographic coordinates system "GCS_NAD_1983_2011" and converted to a shapefile.

As a last step, the planned total spray area (ha) within an application was calculated to assist in the mixing of the herbicides. Selected squares from the one-meter aggregated grid were summed, units covered to hectares and exported as an excel file.

CapMaps (Version 1.04.00) (Capstan Ag Systems, Inc., Topeka, KS) software was used to convert the ModelBuilder shapefile output into a prescription map designating treatment areas where the spray nozzle would receive an ON signal. Areas where no weeds were detected spray nozzles were by default turned OFF. The file was exported in a DAT format as required by the CapView display on the field equipment (Figure 4.6).
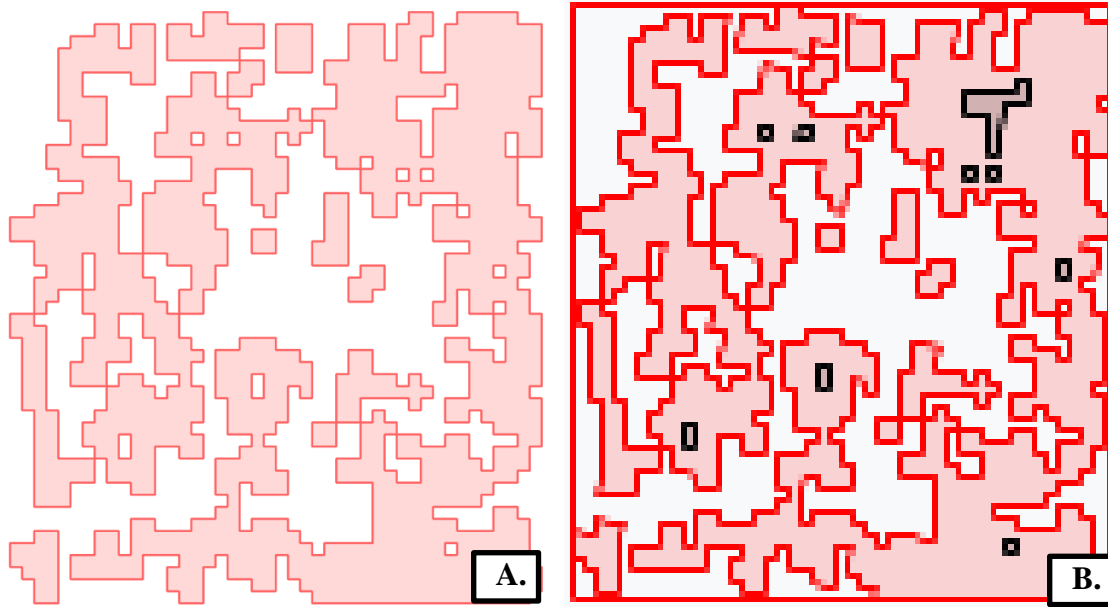
Figure 4.6: Example Modelbuilder output map exported as a shapefile to CapMaps software (A). CapMaps output exported as a DAT map to CapView display (B).

4.3.3   Field application equipment

A high-clearance sprayer (Bowman Mudmaster, Bowman Manufacturing Co., Inc., Newport, AR, USA) was used to apply treatments to all strip plots at 6.44 km/h (Figure 4.7). Treatments were applied using a boom with seven 11002 AIXR flat-fan nozzles (Teejet Technologies, Springfield, IL 62703) on 51-cm spacing. The boom was equipped with a PinPoint II system (Capstan Ag Systems, Inc., Topeka, KS). The PinPoint II PWM system contains high-speed individually controlled spray valves which allow for individual nozzle control. A 20.1 LPM Remco ProFlow 5500 12 VDC (Remco Industries, 10425 Wildlife Loop SE, Osakis, MN 56360) diaphragm pump with an internal 414 kPa pressure switch was used to provide flow and pressure to the site-specific boom. A Hamilton, nylon, 0-1034 kPa pressure regulating bypass valve (WLH CO. LLC, 325 Cherry St, Marshall, MN 49069) was used to provide bypass flow from the boom, enabling manual control of flow rate and therefore calibration. Both broadcast

and site-specific treatments were applied with PWM option on the PinPoint II system disengaged. The broadcast treatment was applied using the manually calibrated boom and all nozzles switched on as the sprayer traversed plots.

The site-specific application enabled 0, or 2-7 nozzles to be on at any given time producing slight variations in system flow, pressure, and resulting rate. These variations were minimized with the mechanical spring bypass built into the Hamilton valve (Figure 4.8). Additionally, since site-specific plots were mapped in a one-meter grid and nozzles were spaced on 51 cm, anytime a weed detection triggered nozzles to spray, at least two and most often three nozzles were automatically engaged depending on how they aligned with the grid. Therefore, the site-specific treatment only varied between two and seven engaged nozzles. The pressure, flow, and rate variation in this range were recorded and are reported with their respective droplet size and ASABE.

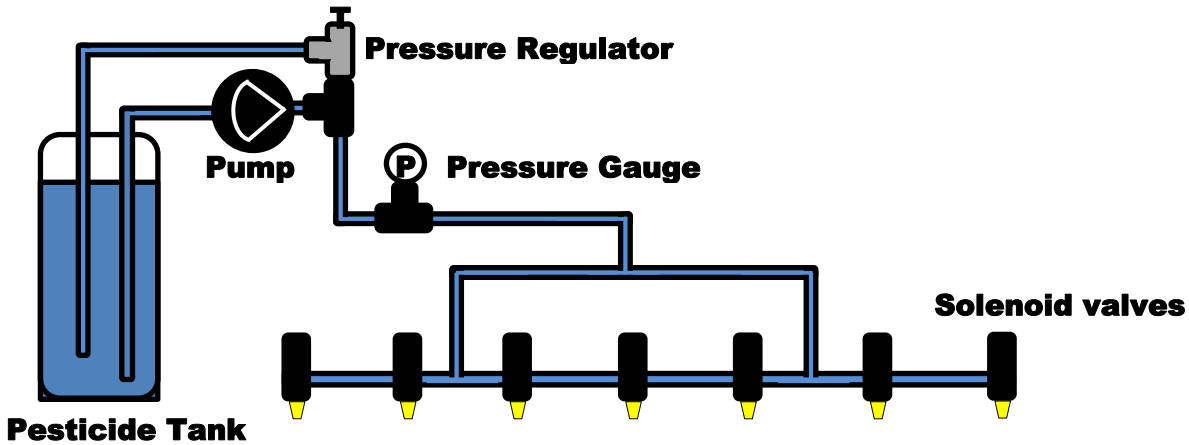**Figure 4.7:** Site-specific Boom mounted on high-clearance sprayer.

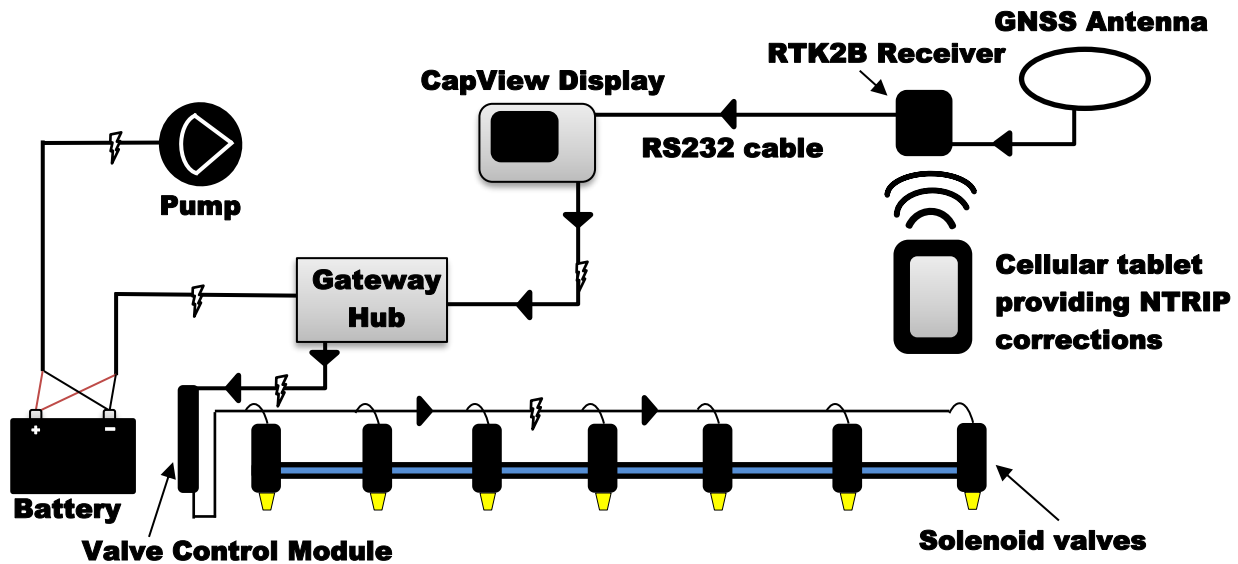**Figure 4.8**: Simplified plumbing diagram of site-specific boom on field equipment.

Figure 4.9: Fundamental electrical components of site-specific boom on field equipment.

The PinPoint II PWM control system was provided real-time positioning via a dual-band RTK2B GNSS receiver (Figure 4.9). The RTK2B GNSS rover leveraged a U-blox ZED-F9P microchip and a survey grade IP66 multiband antenna. Real-time kinematic (RTK) corrections were received through the CORS network maintained by ARDOT and broadcast through an NTRIP caster. The CORS base station used for corrections were from the ARDOT District Five headquarters outside of Batesville, Arkansas, USA approximately 39 km away from the study sites.

PinPoint II system parameters were adjusted to facilitate the site-specific application. This primarily involved adjusting system "look ahead time" to ensure that nozzles were actuated on at the proper time relative to equipment speed. The adjusted PinPoint II system parameters are shown in Table 4.2. These parameters were determined through a series of controlled passes over test detections. Water sensitive cards were placed in the center of test detection boxes, the site-specific boom was enabled, and sprayer was driven at the calibrated 6.4 km/h. The water

sensitive cards acted as indicators of how well settings optimized hitting the target. Look ahead times were tested at 0.0, 0.1, 0.2, and 0.3 seconds. It was determined that at the selected treatment speed (6.4 km/h), and with the 0.5 m radius buffer, no look ahead time was needed (0.0 seconds).

Table 4.2: Applicable parameters selected for site-specific application using the PinPoint II controller.

Line	Parameter	Setting	Why
1	Operation Mode	SharpShooter	Enables x, y, and z
16	Overlap Distance	60 in.	Prevents adjacent nozzles from marking untreated 1 m squares as treated
17	Look ahead time	0 s	Prevents nozzles from turning on prior to entering a treatment zone
18	Zero rate shutoff	Shutoff	No electronic flow meter on system

4.3.4   Field data collection

Calculating treatment areas

Areas where herbicides were applied or treatment areas (m^2) per plot were used to evaluate null hypothesis five. Treatment area (m^2) for broadcast plots were the total plot area. Likewise, treatment area for untreated plots was set to zero. Treatment area per plot for site-specific plots was determined by clipping a copy of the output treatment map (Shapefile for CapMaps in GCS) by the site-specific plot layer. The clipped polygons were summarized by plot area in square meters and exported into an excel file for analysis.

Weed count and size

In year one and two, individual weed counts were determined by surveying the high resolution orthomosaics. Plot specific weed counts for year one and year two within each plot for each week were combined into a single excel file for statistical analysis. Weed size was

111

estimated using the length of side (cm) of the detection boxes. In year two, individual weeds were geolocated and heights (cm) estimated through *in situ* scouting. A linear regression was calculated on the relationship between the *in situ* measurements of weed height to the length of side of detected bounding boxes (Equation 4.1). Bounding box sizes were used to estimate any differences in weed size between treatments within each year, while the measured and calculated weed heights were only used for discussion in the context of herbicide treatment timing relative to weed size.

Equation 4.1 *Linear regression used to describe the relationship between year two's measured values and bounding box size of detected weeds.*

$$y = 0.254x + 0.72x$$

Treatment Efficacy

Pretreatment in-plot weed count were determined by manually counting weeds in the pretreatment imagery. After herbicide treatments were applied, aerial imagery was taken at one-, two-, and three-weeks after treatment (WAT). Weed counts were performed by manually annotating weeds in each orthomosaic (one-, two-, and three-WAT). In year one, individual weed ratings were determined by surveying the high resolution orthomosaics overlayed with the keyed weed map used in comparing model accuracy. Individual weed control ratings were on a scale from 0-10 with 0 being no herbicide symptomology (no control) and 10 representing dead or necrotic weeds (full control). Year two, in addition to the 0.4 cm spatial resolution RGB imagery collected one-, two-, and three-WAT; weed locations, size, and vigor ratings for each weed were manually collected through *in situ* scouting on imaging dates. Figures 4.10, 4.11, and 4.12 show example GNSS points collected manually in the field for each treatment (UTC,

Broadcast, and Site-specific, respectively). A, B, C, and Ds within each figure represent Pre, and

one, two-, and three-weeks post treatment, respectively.

Weed specific ratings for year one and year two within each plot for each week were

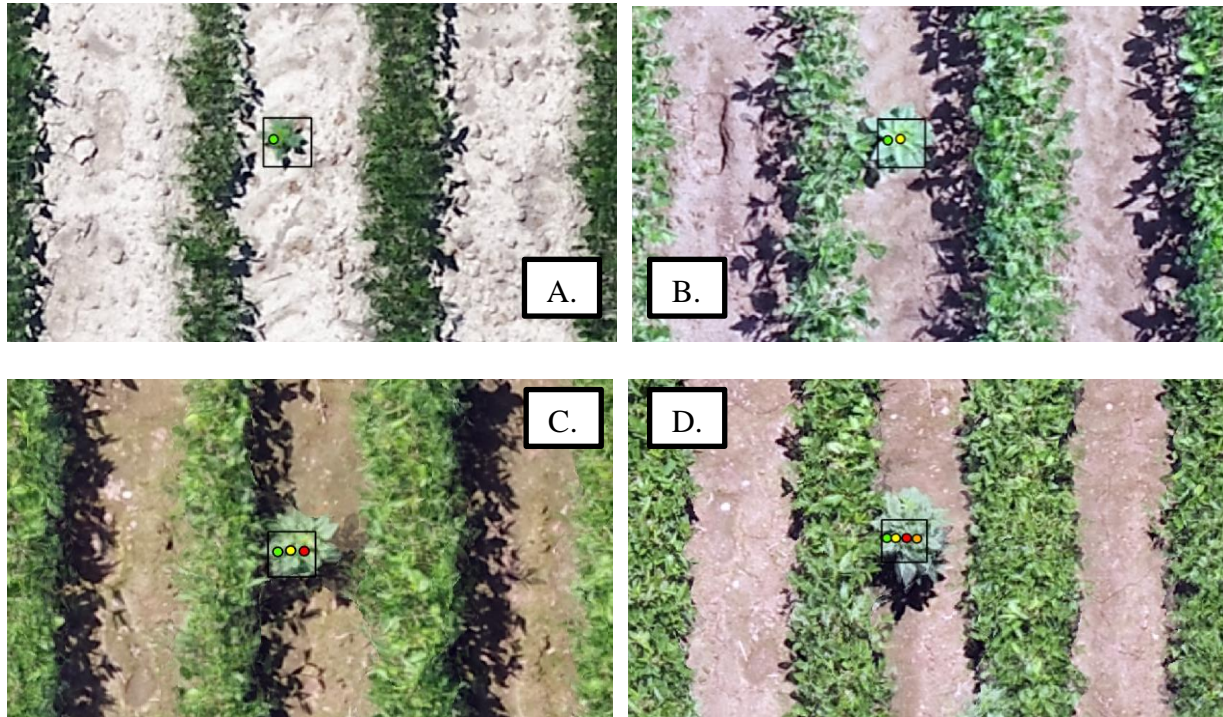combined into a single excel file for statistical analysis.

Figure 4.10: Georeferenced weed ratings in year two -UTC. Pre-treatment scout (A), one WAT (B), two WAT (C), three WAT (D).
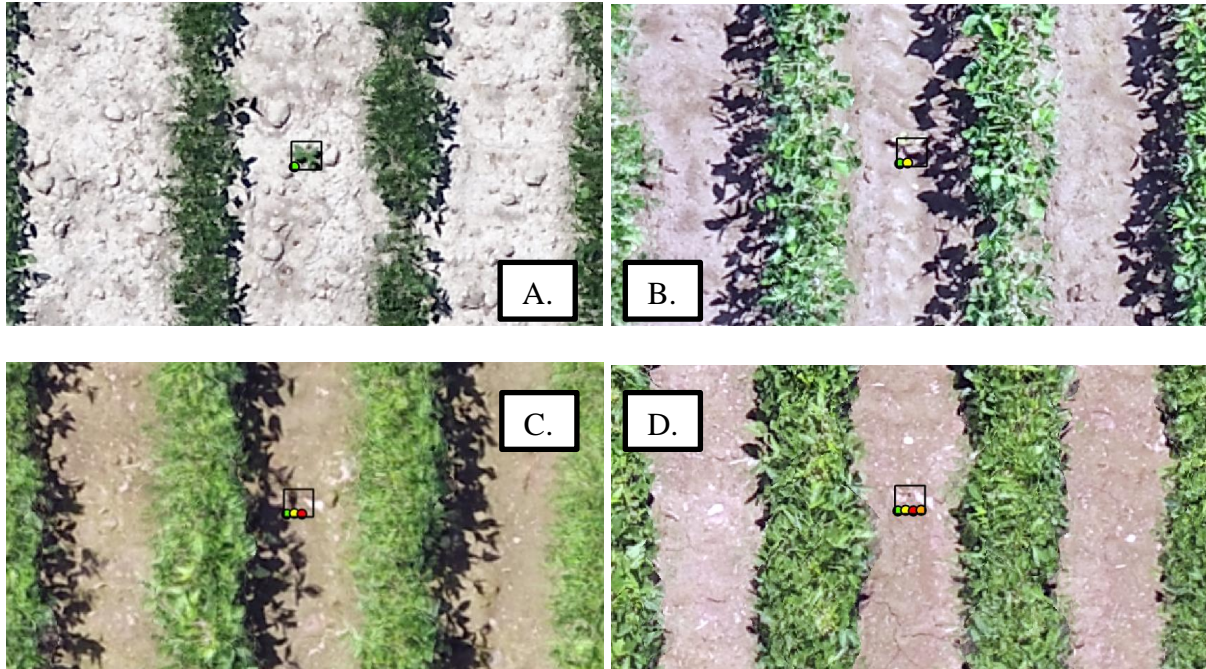
Figure 4.11: Georeferenced weed ratings in year two -Broadcast. Pre-treatment scout (A), one WAT (B), two WAT (C), and three WAT (D).
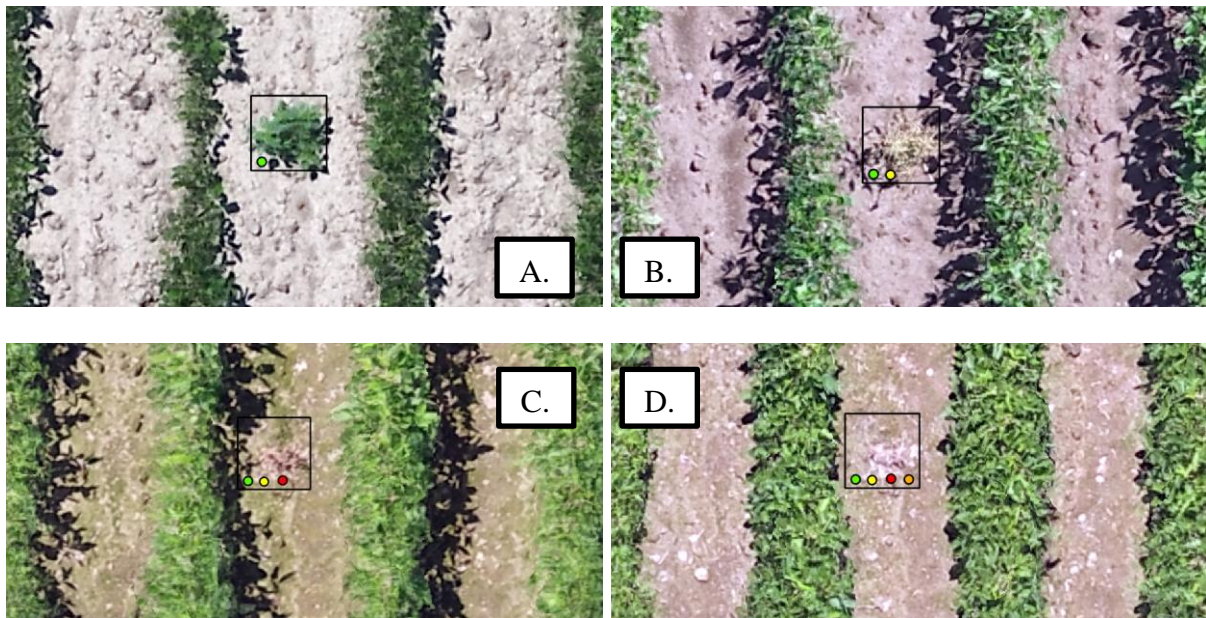

Figure 4.12: Georeferenced weed ratings in year two - Site-specific. Pre-treatment scout (A), one WAT (B), two WAT (C), three WAT (D).

Water Sensitive Cards

SpotOn ® water sensitive papers (Innoquest inc., Woodstock, IL) were placed within plots immediately prior to applying treatments and recovered immediately after treatment was applied. Two cards were randomly placed in each broadcast plot, two cards were placed in known treatment areas within site-specific plots, and two additional cards were placed in known non-treatment areas of site-specific plots. Water sensitive cards were analyzed for spray coverage (%) data using the Java plugin Water Paper Analysis on the desktop software DepositionScan distributed by the United States Department of Agriculture (USDA).

4.3.5   Data analysis

Weed counts, weed size data, treatment ratings (% control), and mean plot spray coverage (%) were averaged to obtain the mean and standard deviation for each plot. Treated area per plot (m^2) was summarized from the application maps, added to the dataset, and used to calculate weed density per plot (weed/m^2).

All data means were entered into RStudios (Version 4.2.1, Bostin, MA) for analysis. A t-test using the "rstatix" package (Version 0.7.0) in RStudio to determine any significant difference ($p < .05$) in weed sizes between year one and year two trials. One-way Analysis of Variance (ANOVA) using the "rstatix" package (Version 0.7.0) was used in RStudio to determine any significant difference ($p < .05$) in weed size and weed density between treatments within each year, and across both years. ANOVA was used to identify any significant difference ($p < .05$) in one, two, and three WAT ratings between treatments within each year and across years. ANOVA was also used to identify any significant differences ($p < .05$) in percent coverage (%) between broadcast, site-specific treated, and site-specific untreated spray cards.

Significant ANOVA values from all sets of means were then subjected to Tukey's post hoc analysis to determine the amount of significance.

4.4   Results

The following section summarizes the quality of orthomosaics, CNN model accuracy, and equipment configuration used to evaluate the proposed site-specific workflow. Treatment areas are described including weed sizes and weed densities. Finally, the efficacy results of field trials comparing broadcast, site-specific, and un-treated check plot work are summarized.

4.4.1   Orthomosaic quality

Table 4.3 summarizes year one and year two trial orthomosaics used for site-specific map creation. While the target spatial resolution during collection events was 0.4 cm, the resulting maps are reported as 0.38 and 0.39 cm spatial resolution for year one and year two maps respectively. Measured from the provided GCPs, orthomosaic RMS error was reported as 1.20 cm and 2.00 cm for year one and two respectively. X and Y errors were less than 2.25 cm and z errors were less than 3.75 cm for both years. The accuracy of individual GCPs for year one is reported in Table 4.4 and illustrated in Figure 4.13. The accuracy of imported GSPs used for correction were all 0.020 m for X, Y, and Z axes. RMS error of all GCPs was reported as 0.013, 0.021, and 0.004 m for the X, Y, and Z axes respectively.

Table 4.3: Accuracy metrics of orthomosaic output from Pix4DMapper on year one and two from quality reports.

Yr.	Flight	Date	Number of images	Calibrated images	Ortho GSD (cm)	Ortho RMS error (cm)	GCP Error X (cm)	GCP Error Y (cm)	GCP Error Z (cm)
1	Pre	21-07-25	278	278	0.38	1.20	1.31	2.07	0.35
2	Pre	22-07-22	264	264	0.39	2.00	1.14	1.89	3.56

Table 4.4: Localization accuracy of year one orthomosaic per GPC's and mean errors in three coordinate directions. Accuracies are derived from quality report generated from Pix4D processing.

| GCP Name	Accuracy XY/Z (m)	Error X (m)	Error Y (m)	Error Z (m)	Projection Error (pixel)	Verified/ Marked Images
SW Corner	0.020/0.020	-0.017	0.035	-0.005	2.203	12/12
NW Corner	0.020/0.020	-0.011	-0.034	0.003	1.721	12/12
NE Corner	0.020/0.020	0.002	0.010	-0.004	1.495	20/20
SE Corner	0.020/0.020	-0.007	-0.003	0.003	1.421	13/13
S Middle	0.020/0.020	0.017	-0.011	0.003	1.595	14/14
N Middle	0.020/0.020	0.017	0.003	0.001	0.931	15/15
Mean (m)		0.000	0.000	0.000		
Sigma (m)		0.013	0.021	0.004		
RMS Error (m)		0.013	0.021	0.004		
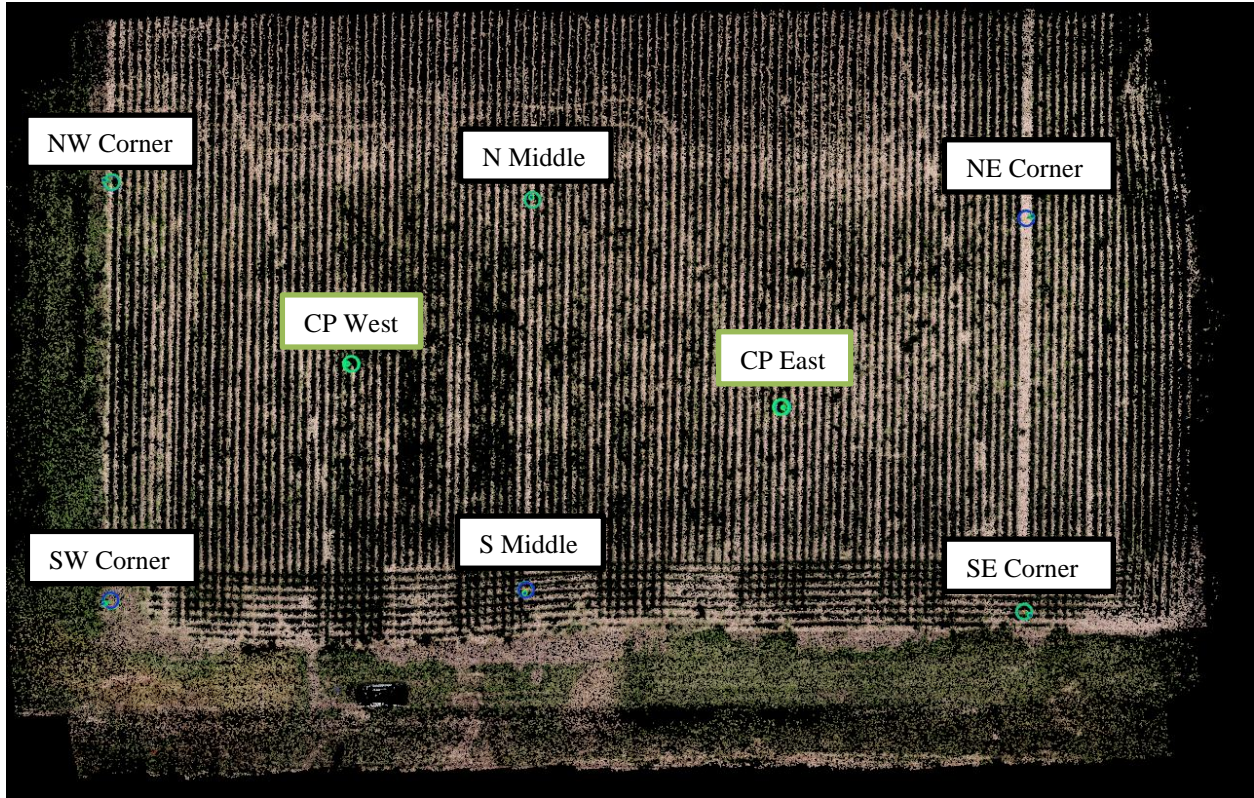
Figure 4.13: Locations of GCPs and CPs used in processing and assessing orthomosaic geolocation accuracy for year one.

Independent verification of orthomosaic geolocation accuracy was done through two CPs that were used for assessment only and not introduced into processing. The reported geolocation accuracy to the CPs is reported in Table 4.5. The RMS error of year one orthomosaic measured by CPs was 0.0968, 0.0043, and 0.0531 m for the X, Y, and Z axes respectively. Reported X and Y accuracies provided confidence that objects detected in the year one orthomosaic map were within 9.68 cm of their actual location in the field.

Table 4.5: Localization accuracy of year one orthomosaic per CPs and mean errors in three coordinate directions. Accuracies are derived from quality report generated from Pix4D processing.

CP Name	Error X (m)	Error Y (m)	Error Z (m)	Projection Error (pixel)	Verified/Marked Images
CP East	0.0277	0.0050	-0.0676	2.3497	17/17
CP West	0.1341	0.0034	-0.0328	2.3104	14/14
Mean (m)	0.0809	0.0042	-0.0502		
Sigma (m)	0.0532	0.0008	0.0174		
RMS Error (m)	0.0968	0.0043	0.0531		

The accuracy of individual GCPs for year two are reported in Table 4.6 and illustrated in Figure 4.14. The accuracy of imported GSPs used for correction were all 0.020 m for X, Y, and Z axes. RMS error of all GCPs was reported as 0.011, 0.019, and 0.032 m for the X, Y, and Z axes respectively.

Table 4.6: Localization accuracy of year one orthomosaic per GPC's and mean errors in three coordinate directions. Accuracies are derived from quality report generated from Pix4D processing.

GCP Name	Accuracy XY/Z (m)	Error X (m)	Error Y (m)	Error Z (m)	Projection Error (pixel)	Verified/ Marked Images
NW Corner	0.020/0.020	-0.008	-0.026	0.015	1.223	20/20
N 3	0.020/0.020	0.008	-0.009	0.036	0.816	20/20
NE Corner	0.020/0.020	-0.002	0.022	-0.016	0.610	19/19
M1 3	0.020/0.020	0.016	-0.009	0.036	0.794	15/15
M1 2	0.020/0.020	0.001	-0.004	-0.046	1.334	17/17
M1 1	0.020/0.020	0.004	0.029	-0.052	0.751	15/15
M2 1	0.020/0.020	-0.018	0.011	0.025	0.402	15/15
M2 3	0.020/0.020	0.006	-0.020	0.029	1.153	15/15
M2 4	0.020/0.020	-0.023	0.006	-0.034	1.118	15/15
S 2	0.020/0.020	0.000	0.018	0.025	0.618	15/15
S 3	0.020/0.020	0.015	-0.032	-0.024	1.076	19/19
SE Corner	0.020/0.020	-0.001	0.017	0.010	1.131	20/20
Mean (m)		-0.000	0.000	0.000		
Sigma (m)		0.011	0.019	0.032		
RMS Error (m)		0.011	0.019	0.032		

Independent verification of geolocation was done through four CPs that were only used in assessment and not introduced into processing. The reported geolocation accuracy to the CPs is reported in Table 4.7. The RMS error of year two orthomosaic measured by CPs was 0.013, 0.010, and 0.070 m for the X, Y, and Z axes respectively. Reported X and Y accuracies provided confidence that objects detected in the year two orthomosaic map were within 1.2 cm of their actual location in the field.
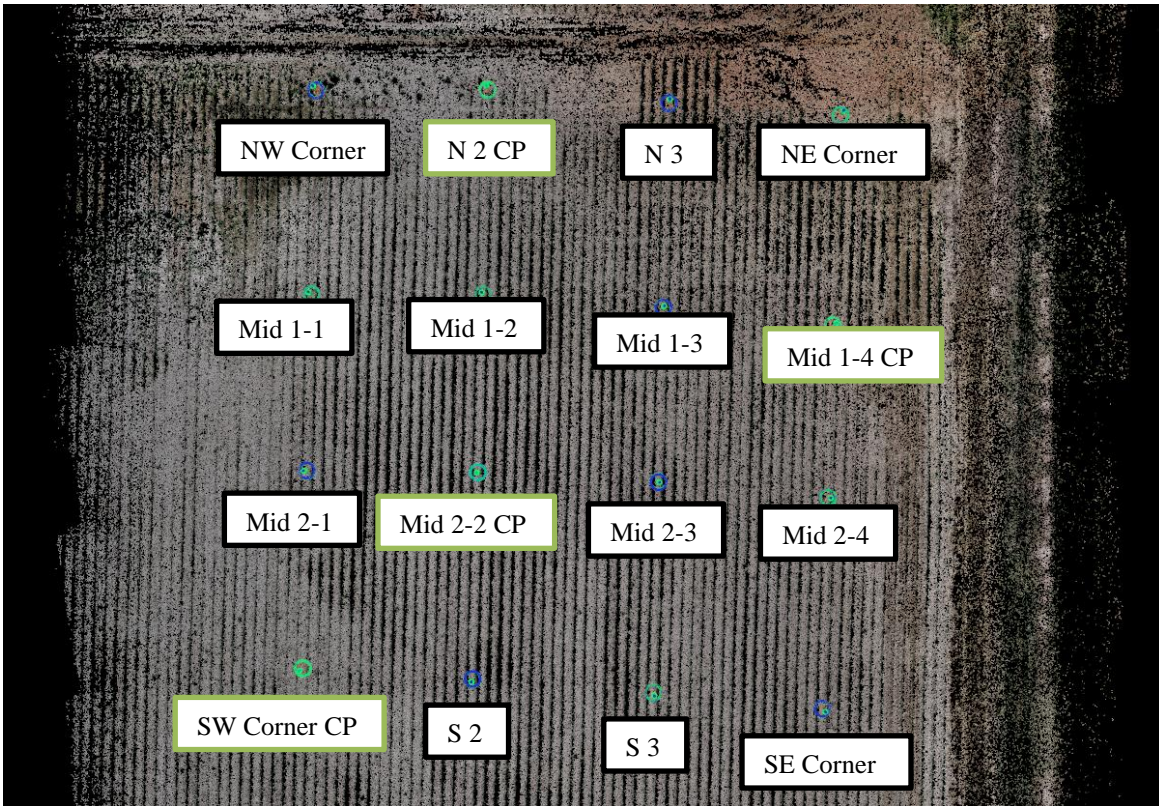
Figure 4.14: Locations of GCPs and CPs used in processing and assessing orthomosaic geolocation accuracy for year two.

Table 4.7: Localization accuracy of year one orthomosaic per CPs and mean errors in three coordinate directions. Accuracies are derived from quality report generated from Pix4D processing.

CP Name	Error X (m)	Error Y (m)	Error Z (m)	Projection Error (pixel)	Verified/Marked Images
N 2	-0.0024	-0.0173	-0.0412	0.8351	18/18
M1 4	-00193	0.0093	-0.0323	0.6483	15/15
M2 2	0.0019	0.0026	0.0433	0.9631	15/15
SW Corner	-0.0170	0.0052	-0.126	2.9648	16/16
Mean (m)	-0.009187	-0.000045	-0.038193		
Sigma (m)	0.009116	0.010222	0.058731		
RMS Error (m)	0.012942	0.010222	0.070057		

4.4.2 Model Accuracies

In year one, a custom trained YOLO v3 CNN model trained on 0.4 cm spatial resolution imagery was used for field testing. Its accuracy was assessed on the field imagery and reported in Table 4.8. It was chosen for field testing due to its recall results of 0.90 which were the highest of models tested at the time of year one trials. The year one YOLO v3 model also reported a precision of 0.58 and F1 Score of 0.70.

Table 4.8: Field specific CNN, GSD, and model accuracy for year one and two.

Year	CNN	Imagery GSD (cm)	Precision	Recall	F1 Score	True Positive	False Positive	False Negative
1	YOLO v3	0.4	0.58	0.90	0.70	1341	985	145
2	FR-CNN	0.4	0.84	0.90	0.87	332	64	37

In year two, a Faster R-CNN model trained on 0.4 cm GSD imagery was used for field testing. Its accuracy was assessed on year two field imagery and reported in Table 4.8. It was chosen for field testing due to its recall results of 0.90 which were the highest of models tested at the time of year two trials. The year two Faster R-CNN model reported a precision of 0.84 and F1 Score of 0.87.

4.4.3 Sprayer configuration

Measured flow rates (LPM), pressure (kPa), calculated spray volume (L ha^{-1}), resulting rate bump (%) and ASABE droplet classification from the experimental sprayer are reported in Table 4.9. When seven of seven nozzles were engaged, the flow rate and pressure of each measured 0.76 LPM and 276 kPa, resulting in the targeted 139.4 L ha^{-1}. At this pressure, the ASABE droplet classification is listed as "Coarse" in the manufactures catalog (Teejet, 2014).

Nozzle measurements were unchanged when only six and five nozzles were engaged indicating that the mechanical bypass valve accounted for increasing bypass need and rate was held constant. When only three and four nozzles were engaged flow rates from nozzles increased to 0.79 LPA, pressure increased to 310 kPa which resulted in a 4% rate bump to 145.0 L ha^{-1}. When two nozzles were engaged, flow rates from nozzles increased to 0.83 LPA, pressure increased to 345 kPa which resulted in a 9% rate bump to 152.5 L ha^{-1}. Since site-specific maps were on a one-meter grid, it was rare for only two nozzles to be engaged and there were no areas where only a single nozzle was engaged. ASABE droplet classification remained "coarse" through all nozzle configurations.

Table 4.9. Nozzle spray characteristics by number of nozzles spraying site-specific boom.

Nozzles spraying	Nozzle Flow rate (LPM)	Pressure (kPa)	Spray volume (L ha^{-1})	Rate increase	ASABE classification - AIXR 11002 @ reported pressure (Teejet, 2014)
7	0.76	276	139.4	+ 0%	Coarse
6	0.76	276	139.4	+ 0%	Coarse
5	0.76	276	139.4	+ 0%	Coarse
4	0.79	310	145.0	+ 4%	Coarse
3	0.79	310	145.0	+ 4%	Coarse
2	0.83	345	152.5	+ 9%	Coarse

4.4.4 Weed sizes

As with production fields, variations in weed size and maturity existed. An independent t-test was conducted to explore differences between weed size in year one and year two plots. An alpha level of .05 was used (Table 4.10). A statistically significant ($p < .05$) difference was found in the mean bounding box sizes between years, $t(420.49) = 17.00$, $p < .001$.

Table 4.10: Descriptive statistics for bonding box size by trial year.

Bounding box size	N	Mean	SD
Year 1	1323	19.62	7.19
Year 2	223	13.31	4.72

Figure 4.15 summarizes the sizes of bounding boxes representing detections and their distribution in year one trial and year two. Year one trial had a significantly larger bounding box detections compared to year two, indicating larger weed size. This correlated with *in situ* plot observations. This suggests that any differences in treatment efficacies between years, may be influenced by the size of the weeds treated.
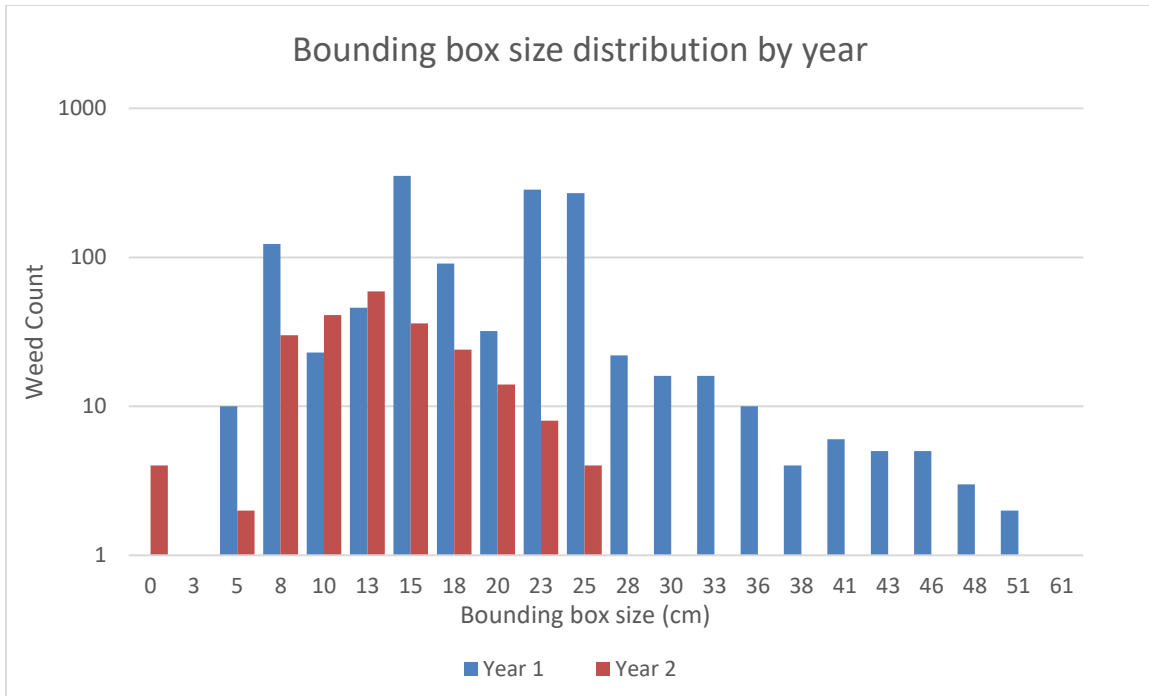
Figure 4.15: Bounding box size distribution by year. Year one trial had a significantly larger bounding box detections compared to year two, correlating to larger weed size.

Mean bounding box size for year one was 19.62 cm (SD = 7.19) while mean bounding box size for year two was significantly smaller at 13.31 cm (SD = 4.72). Year two weed heights were measured in the field (M = 8.99 cm, SD = 4.19). Trial two measured heights and bounding box size were corelated using a least squares approach resulting in an R^2 value of 0.6. The linear relationship describing this correlation is shown in Equation 4.1. Equation 4.1 was used to estimate the weed sizes of year one. Calculated weed height for year one was 14.83 cm (SD = 5.13). Mean measured heights, bounding box lengths, and calculated weed heights are summarized in Figure 4.16.
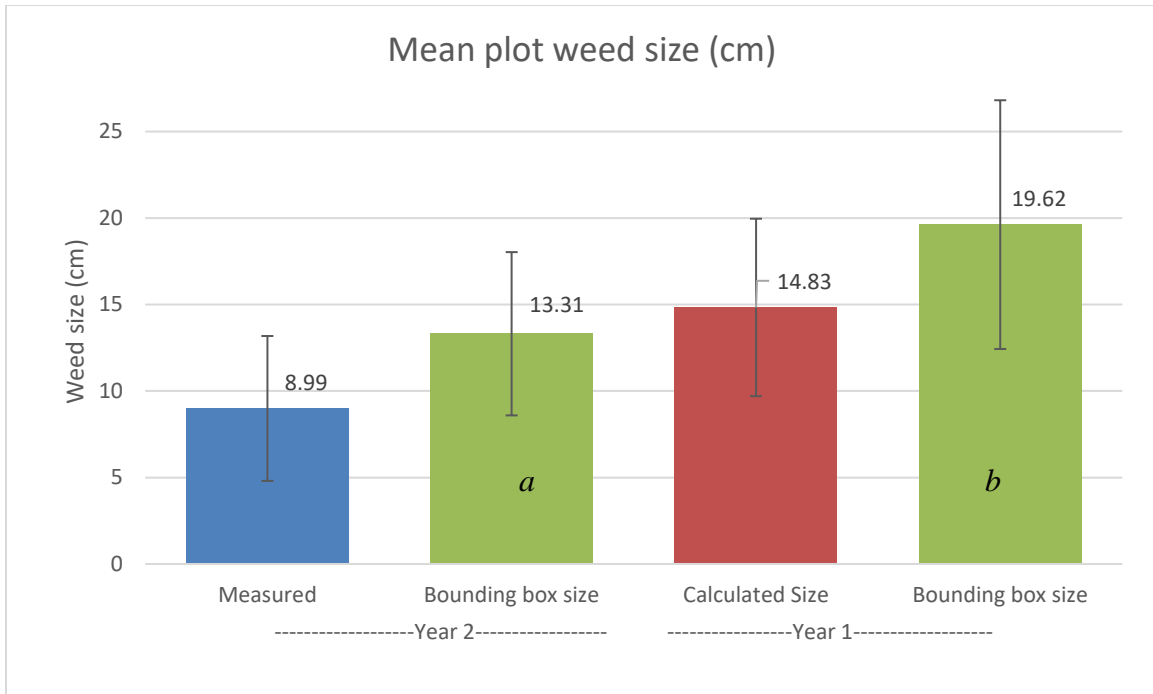
Figure 4.16: Bounding box sizes, measured weed height, and calculated weed height by year. Means of different letters are statistically (p < .05) significant.

Mean plot bounding box sizes by treatment were analyzed using an ANOVA. Table 4.11. shows that ANOVA results for box size where no significant differences were found in mean bounding box sizes between treatments ($p < .05$) within year one, two and across years one and two. Means, and standard deviations for bounding box size between treatments are summarized Table 4.12. The lack of statistical differences in weed sizes between treatment specific plots suggests that any differences in weed control can be attributed to herbicide treatment method.

Table 4.11: ANOVA results for bounding box size between treatments for year one, year two, and year one and two.

Bounding box size	df	F	p
Year 1			
Treatments	2	2.636	.072
Residuals	1320		
Year 2			
Treatments	2	0.122	.885
Residuals	220		
Year 1 and 2			
Treatments	2	1.387	.25
Residuals	1543		

Table 4.12: Means and standard deviations of bounding box size (cm) between treatments within years and across years one and two.

	Broad		SS		UTC	
	M	*SD*	*M*	*SD*	*M*	*SD*
			-----%-----			
Year 1	19.91	6.90	18.98	6.79	20.00	8.28
Year 2	13.11	4.75	13.44	5.21	13.23	4.12
Year 1 and 2	18.91	7.05	18.26	6.86	18.92	8.15

Note. Numbers within rows that share subscript letters are not significantly different ($p \geq .05$)

4.4.5 Weed density

Weed counts along with plot size from years one and two were used to calculate weed density (plants/m^2). Mean weed densities by plot were analyzed using an ANOVA. Table 4.13

shows that ANOVA results for weed densities where no significant ($p < .05$) differences between

treatment means were found within year one, two and across years one and two. Means, and

standard deviations for weed density (weed/m^2) between treatments are summarized in Table

4.14 and Figure 4.17. The lack of statistical ($p \geq .05$) differences between treatment plots within

years and across years increases the confidence that differences found in weed control can be

attributed to herbicide treatment method.

Table 4.13: ANOVA results for weed density between treatments for year one, year two, and year one and two.

Weed density	df	F	p
Year 1			
Treatments	2	1.93	0.188
Residuals	12		
Year 2			
Treatments	2	1.61	0.225
Residuals	20		
Years 1 and 2			
Treatments	2	0.87	0.429
Residuals	35		

Table 4.14: Means and standard deviations of plot weed density (weeds/m^2) between treatments within years and across years one and two.

	Broad		SS		UTC	
	M	*SD*	*M*	*SD*	*M*	*SD*
	---------------------------------------%--					
Year 1	1.07	0.21	0.77	0.26	0.99	0.39
Year 2	0.20	0.12	0.12	0.11	0.10	0.10
Years 1 and 2	0.57	0.47	0.43	0.38	0.34	0.46

Note. Numbers within rows that share subscript letters are not significantly different ($p \geq .05$)
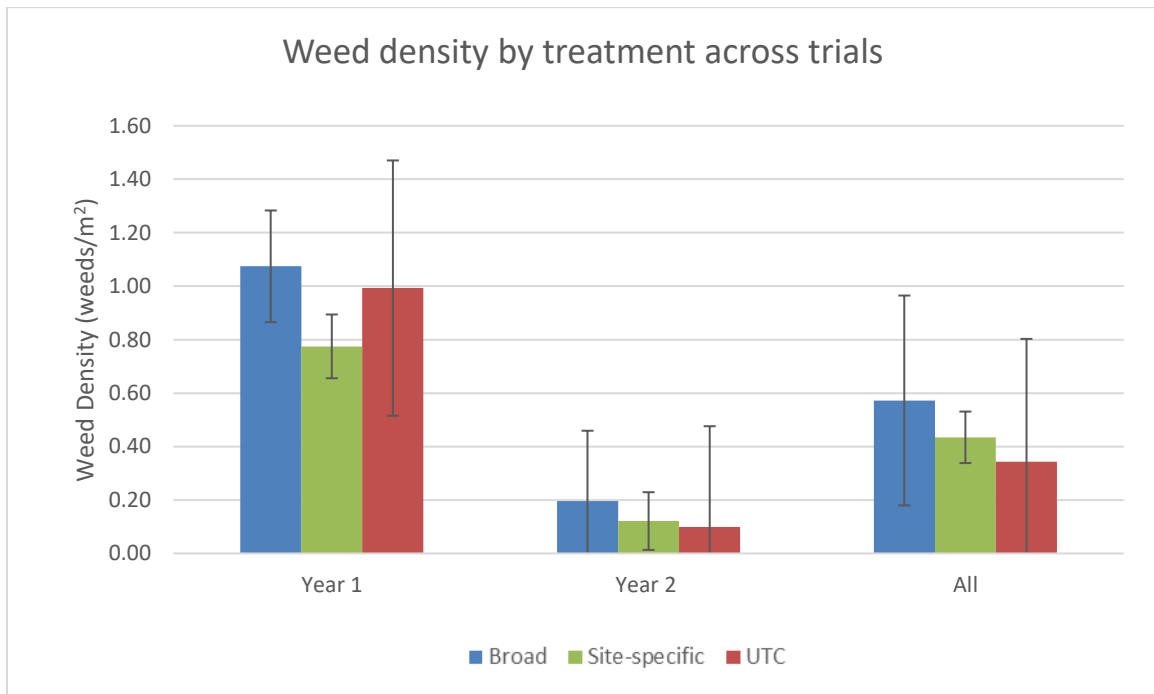
Figure 4.17: Year one and two weed density (weeds/m^2) by treatment.

4.4.6 Comparison of treated areas

Applied area (m^2) for year one and two was summarized by treatment. Mean treatment applied area (m^2) was analyzed using an ANOVA for year one and two. Table 4.15 shows that

ANOVA results for applied area were all significant between treatments ($p < .05$) within year

one, two and between year one and two. Null hypothesis five stated that there will be no

significant ($p < .05$) differences in the mean treatment area (%) between plots sprayed with a

broadcast rate and plots sprayed using prescription map generated by a SSWM workflow.

Therefore, null hypothesis five was rejected. Means, standard deviations, and significant

differences for applied area (m^2) between treatments for each year and across both years are

summarized by Table 4.16 and Figure 4.18.

Table 4.15: ANOVA results for applied area between treatments for year one, year two, and year one and two.

Area treated	df	F	p
Year 1			
Treatments	2	140.9	<0.0001
Residuals	12		
Year 2			
Treatments	2	375.7	<0.0001
Residuals	20		
Years 1 and 2			
Treatments	2	99.4	<0.0001
Residuals	35		

Table 4.16: Percent of Broadcast, site-specific, and untreated check plot areas treated.

	Broad		SS		UTC	
	M	*SD*	*M*	*SD*	*M*	*SD*
	--%--					
Year 1	1.00$_A$	0.00	0.80$_B$	0.13	0.00$_C$	0.00
Year 2	1.00$_A$	0.00	0.27$_B$	0.14	0.00$_C$	0.00
Year 1 and 2	1.00$_A$	0.00	0.52$_B$	0.30	0.00$_C$	0.00

Note. Numbers within rows that share subscript letters are not significantly different ($p \geq .05$).
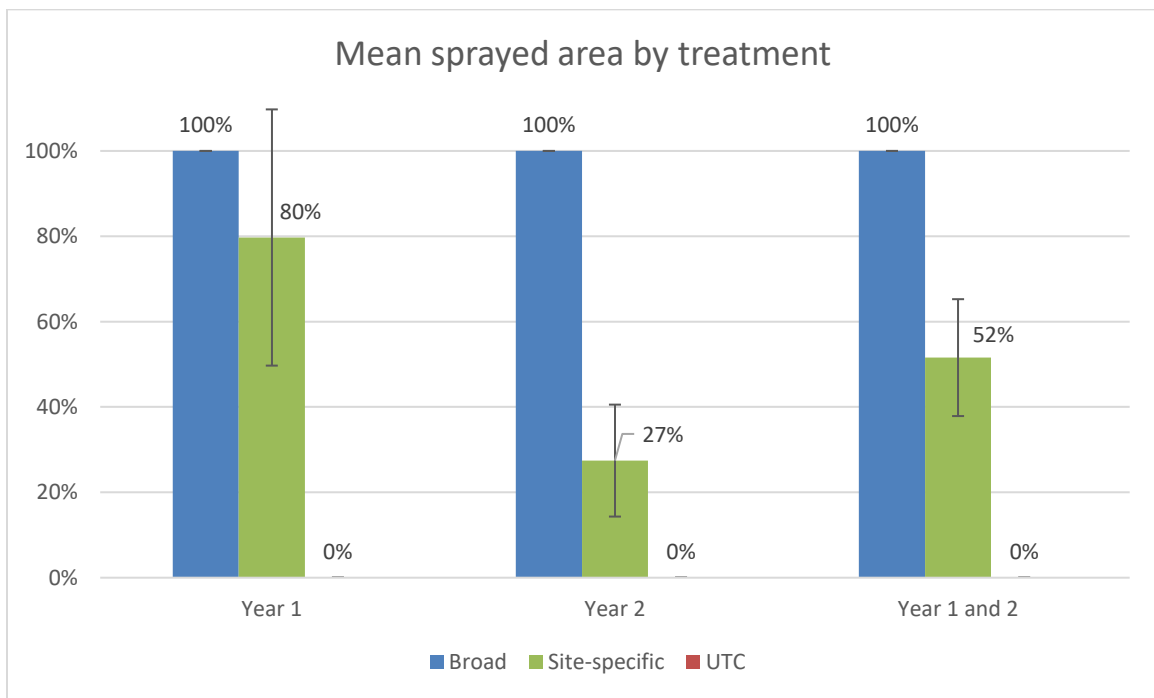
Figure 4.18: Year one, Year two, and combined percent area sprayed averaged by treatment.

There were significant ($p < .05$) differences in mean applied area (%) between broadcast, site-specific and untreated check plots within each year and across both years (Table 4.16 and Figure 4.18). Even with the high weed density in year one, site specific plots had a statistically significant ($p < .05$) 20% reduction in mean applied area. The lower relative weed density in year

two plots increased the difference in applied area to 73%, also significant ($p < .05$). Across years one and two, a 48% reduction in mean applied area was found between broadcast and site-specific plots.

Figure 4.19 displays weed density means with percent plot treated of individual broadcast plots. Year one contained replications one through six and were numerically highest in weed density ranging from 0.83 to 1.34 weeds per m^2. Year two replications seven through 14 had a consistently lower numerical weed density ranging from 0.37 to 0.06 weeds per m^2. As shown, independent of mean weed density, 100% of the broadcast plots were treated.
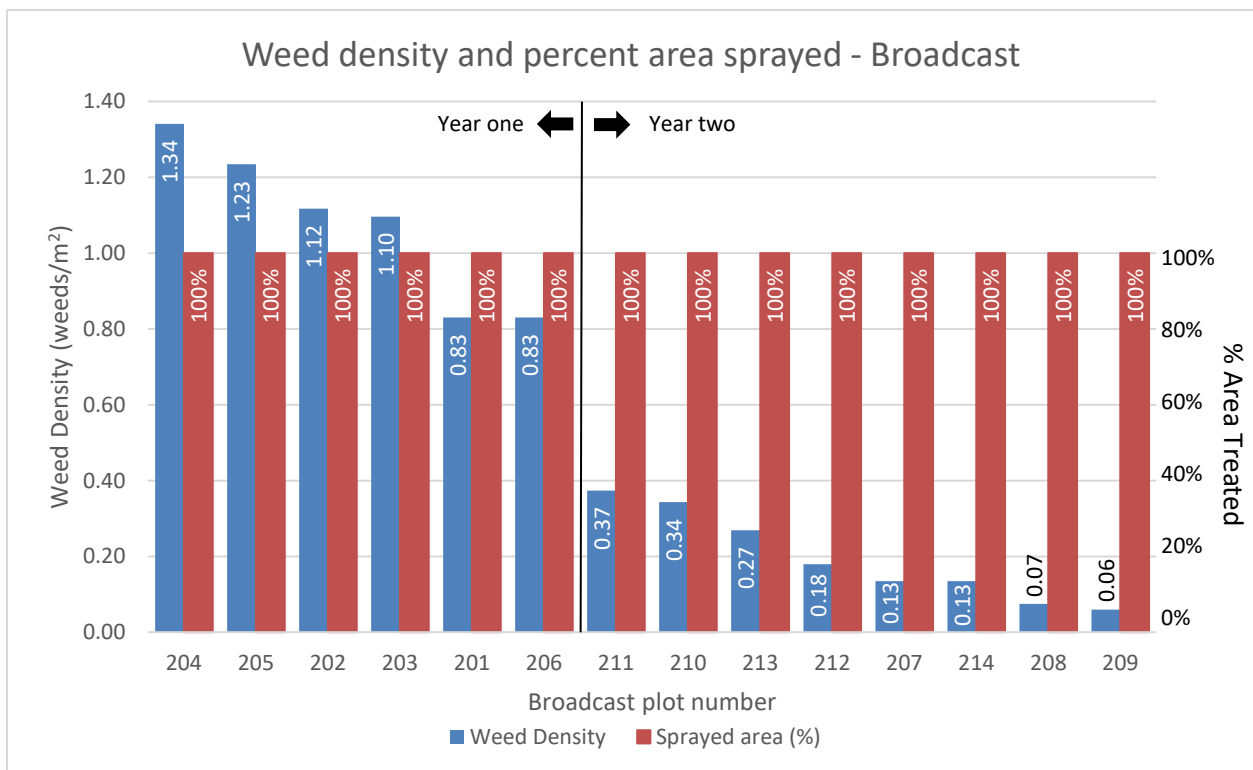
Figure 4.19: Weed density and percent area sprayed of broadcast treatment by plot.

Figure 4.20 summarizes weed density and percent area treated of site-specific plots. Year one contained replications one through six and were ranked highest in weed density ranging from

1.19 to 0.38 weeds per m^2. Year two replications seven through 13 had a consistently lower weed density ranging from 0.27 to 0.00 weeds per m^2. Unlike the broadcast plots, weed density directly influenced applied area and percent treated area ranged between 94% and 7%.
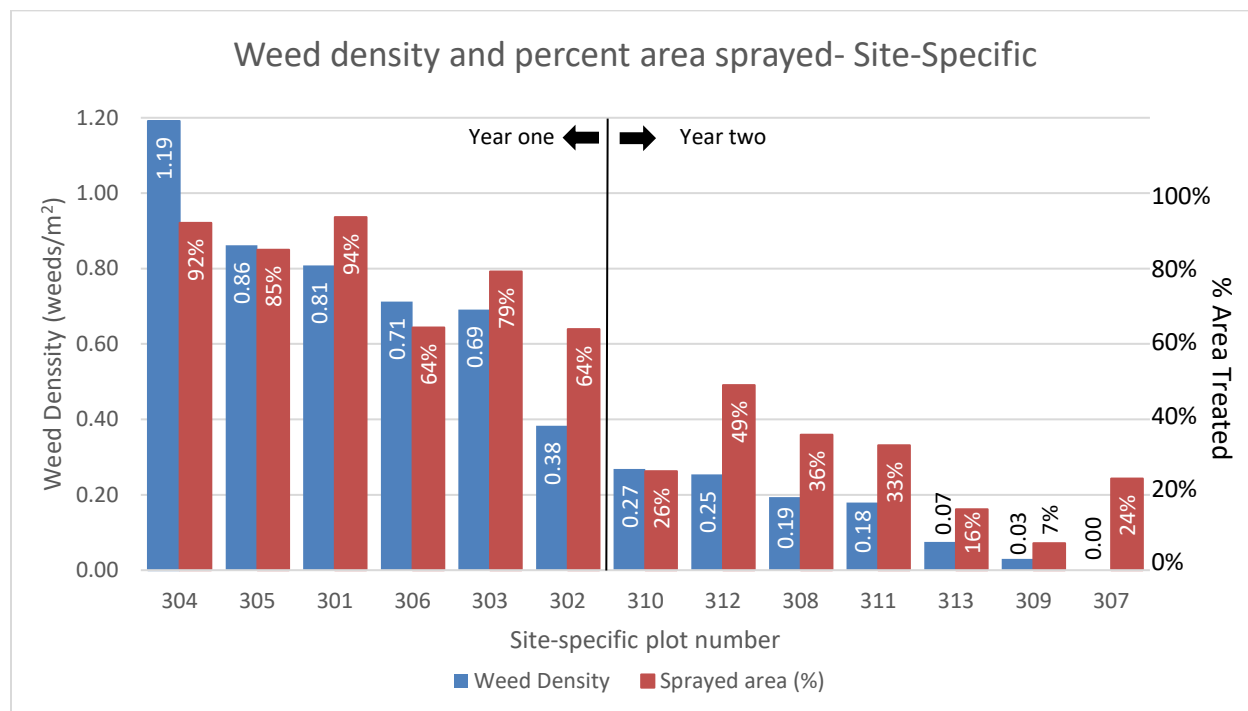

Figure 4.20: Weed density and percent area sprayed of site-specific treatment by plot.

4.4.7   Spray card results

SpotOn ® water sensitive papers (Innoquest inc., Woodstock, IL) collected spray coverage data in year two plots were analyzed for percent coverage and averaged across treatments. Mean treatment coverage (%) were analyzed using an ANOVA. Table 4.17 shows that ANOVA results for percent coverage were significant ($p < .05$). Once ANOVA results were determined to be significant, the means were then subjected to Tukey's post hoc analysis to determine the amount of significance between each treatment. Means, standard deviations, and

significant differences for spray coverage (%) between treatments are summarized Table 4.18
and Figure 4.21.

Table 4.17: ANOVA results for year two water sensitive paper coverage samples.

Coverage (%)	df	F	p
Year two			
Treatments	2	34.3	<0.0001
Residuals	41		

Table 4.18: Percent coverage of Broadcast, site-specific treated, and site-specific non-treated plot areas.

	Broad		SS-Treated		SS-Untreated	
	M	*SD*	*M*	*SD*	*M*	*SD*
Coverage (%)	26.74 a	8.55	28.95 a	11.30	3.51 b	6.93

Note. Numbers within rows that share subscript letters are not significantly different ($p \geq .05$).

There were no significant ($p < .05$) differences found in the mean spray coverage (%) of broadcast areas (M = 26.75%, SD = 8.55) and site-specific areas (M = 28.95%, SD = 11.30) that were designated as treatment zones. There was a significant ($p < .05$) decrease in spray coverage (%) in the areas sampled within site-specific plots that were designated as untreated (M = 3.51%, SD = 6.93) when compared to broadcast and treated areas within site-specific plots (Figure 4.21). Figures 4.22 - 4.25 display the water sensitive cards collected from the treatment areas. Fourteen replications of each test were collected, two spray cards in each treatment designation (broadcast, site-specific treated, site-specific untreated) in each plot. The spray cards clearly demonstrate that similar coverage was achieved in sampled treatment areas between treated areas (broadcast

and site-specific treated). Additionally, areas designated as untreated within site-specific plots by comparison contain little or no coverage.
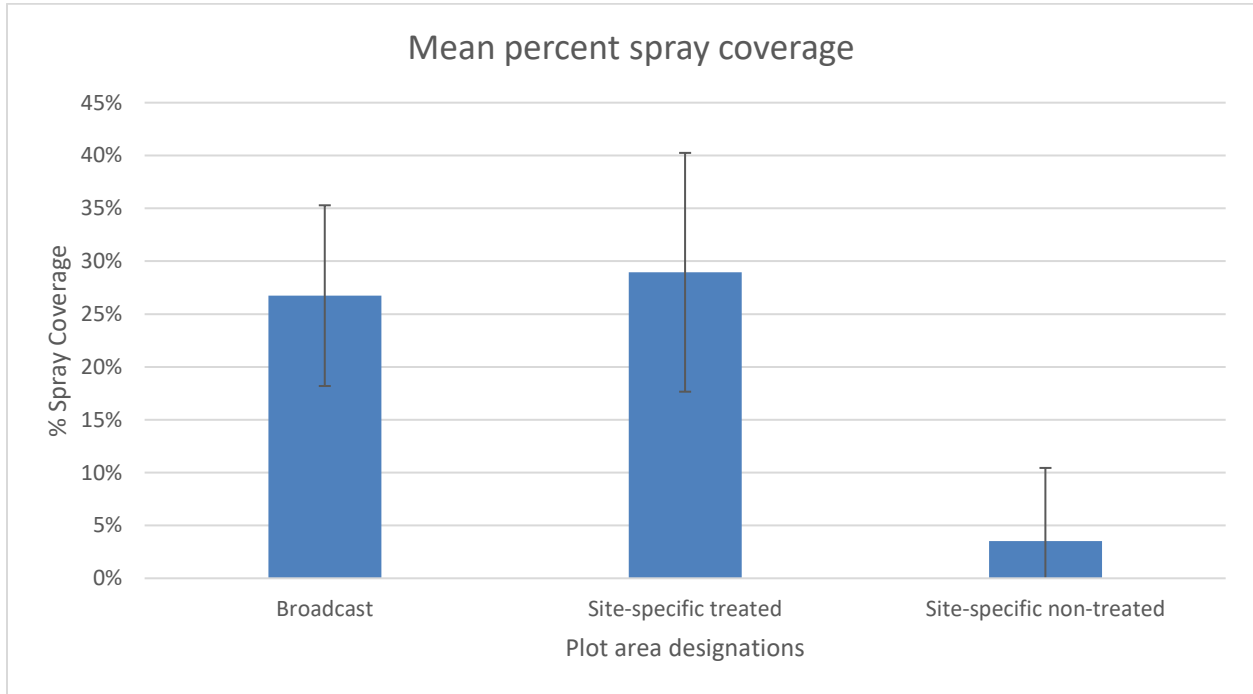

Figure 4.21: Mean spray coverage (%) measured by water sensitive paper within broadcast, site-specific treated, and site-specific untreated areas of plots.
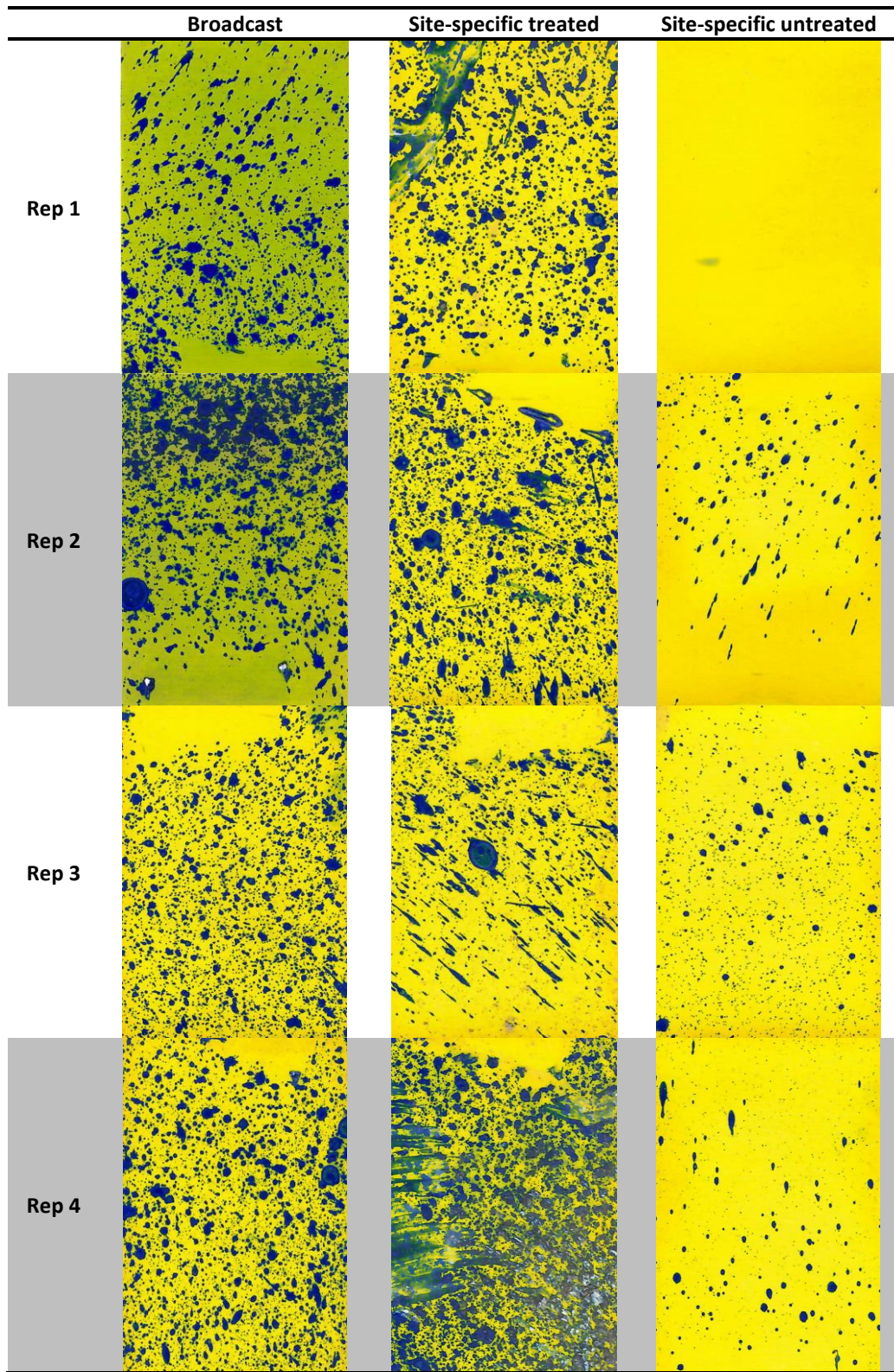
Figure 4.22: Spray coverage illustrated by water sensitive paper within broadcast, site-specific treated, and site-specific untreated areas of plots (Replications 1-4).
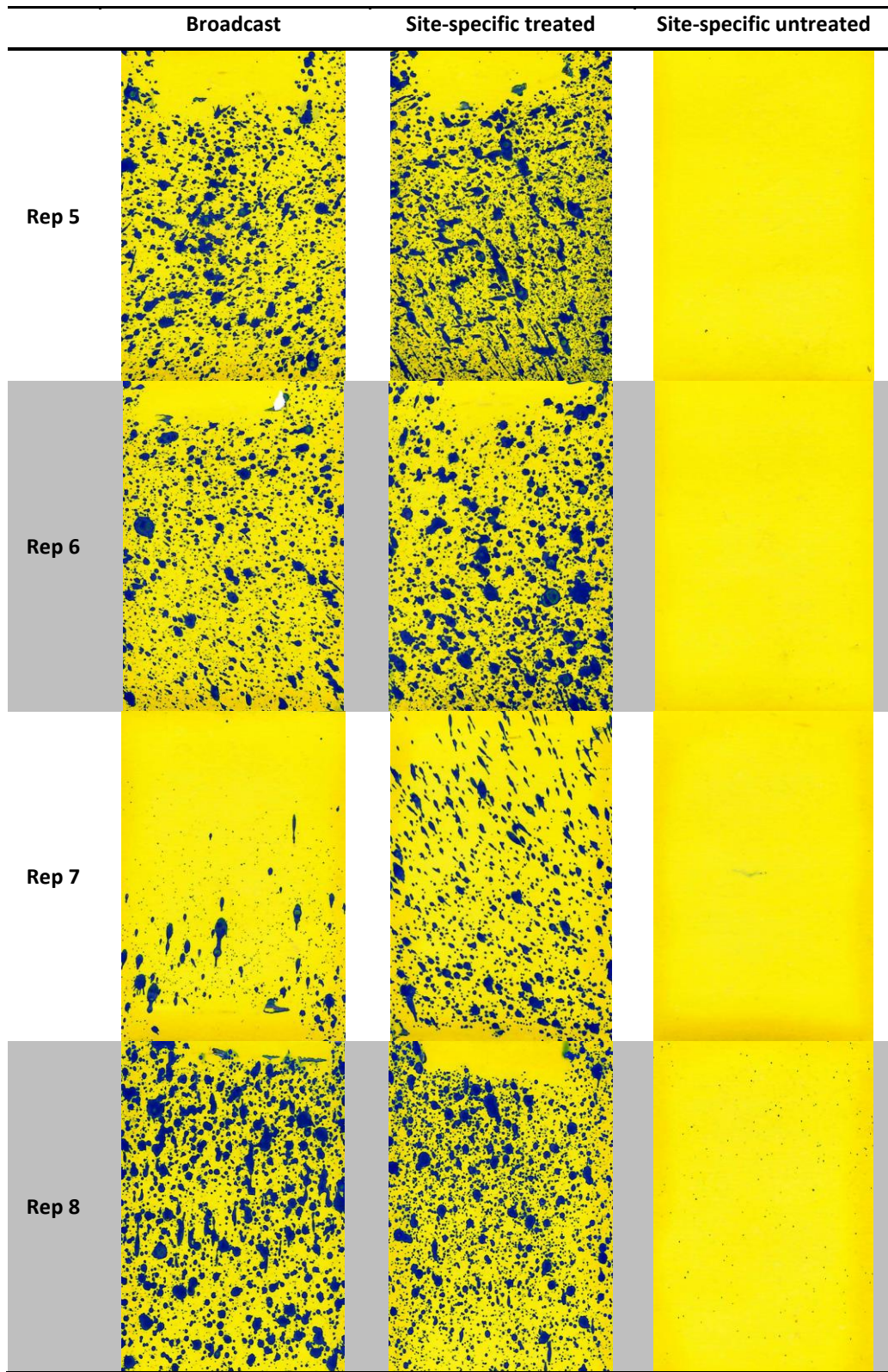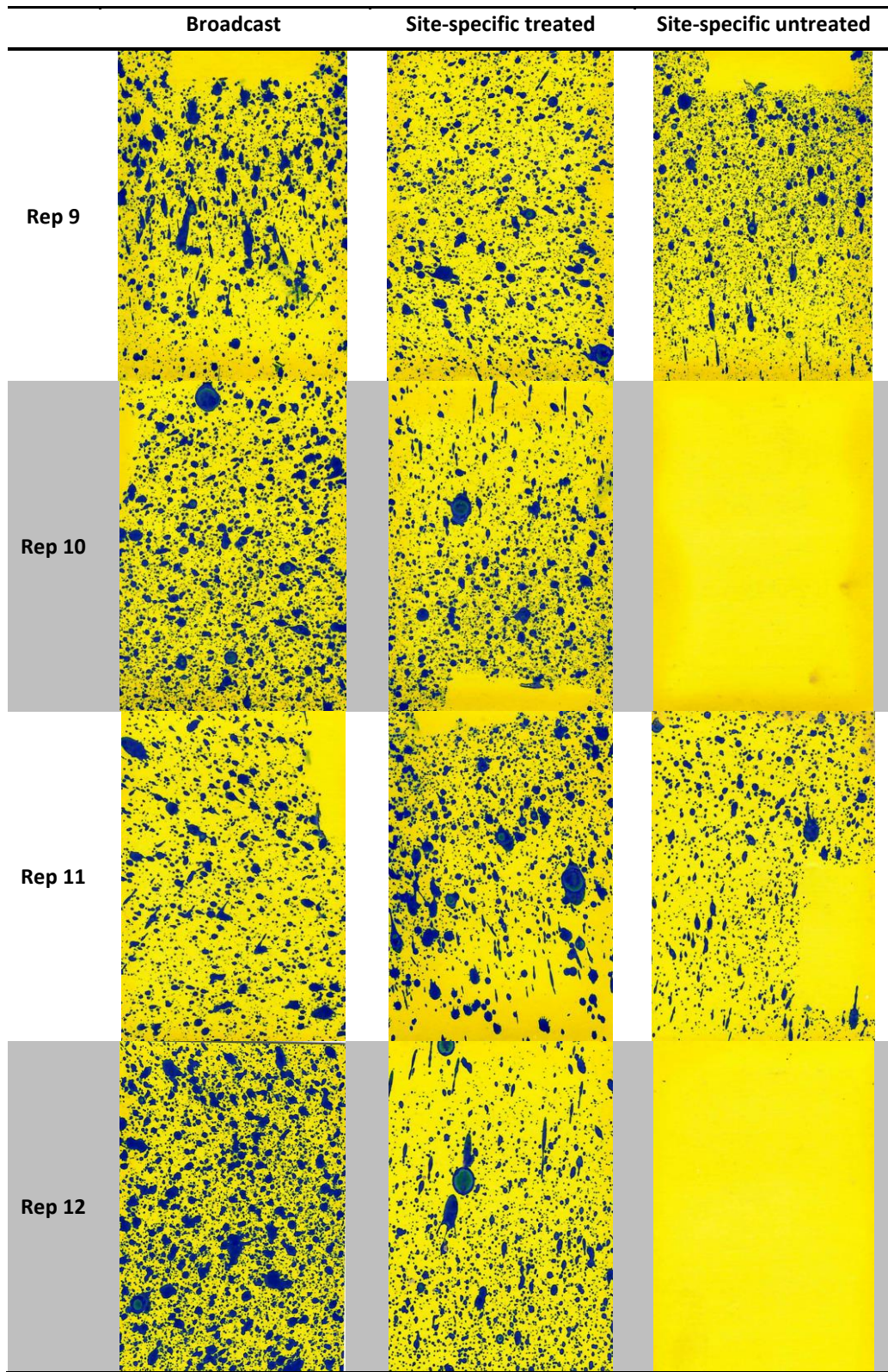
Broadcast	Site-specific treated	Site-specific untreated

Figure 4.23: Spray coverage illustrated by water sensitive paper within broadcast, site-specific treated, and site-specific untreated areas of plots (Replications 5-8).

Figure 4.24: Spray coverage illustrated by water sensitive paper within broadcast, site-specific treated, and site-specific untreated areas of plots (Replications 9-12).
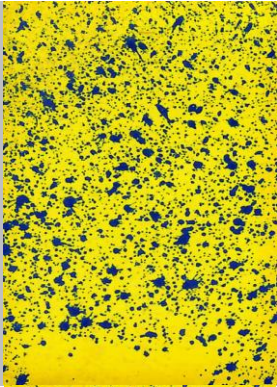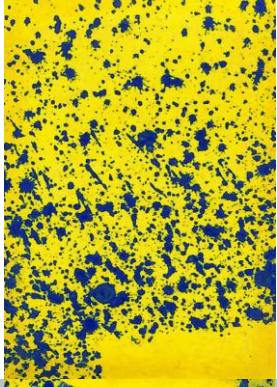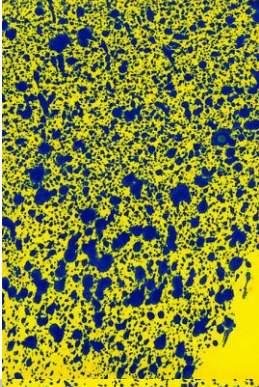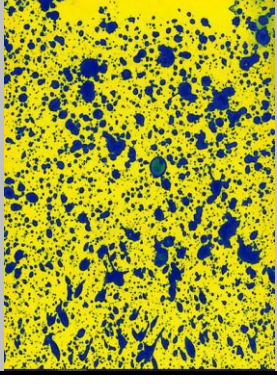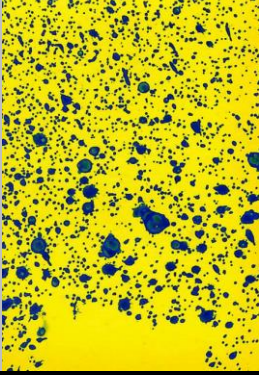
	Broadcast	Site-specific treated	Site-specific untreated
Rep 13		Rep voided, plot missapplication	Rep voided, plot missapplication
Rep 14		Rep voided, plot missapplication	Rep voided, plot missapplication
Rep 15			
Rep 16			

Figure 4.25: Spray coverage illustrated by water sensitive paper within broadcast, site-specific treated, and site-specific untreated areas of plots (Replications 13-16).

4.4.8   Treatment Efficacy

Treatment Efficacy Year One

Individual weed control ratings at one, two and three weeks after treatment for year one was averaged by plot. Mean weed control (%) was analyzed using a One-way ANOVA for individual weeks (one, two, and three weeks after treatment) in year one. Table 4.19 shows that ANOVA results for weed control were all significant between treatments ($p < .05$) within weeks in year one. Once ANOVA results were determined to be significant, the means were then subjected to Tukey's post hoc analysis to determine the amount of significance between each treatment. Means, standard deviations, and significant differences are for year one weed control (%) between treatments are summarized in Table 4.20 and Figure 4.26.

Table 4.19: ANOVA results for year one treatment efficacy.

Year one	df	F	p
1 WAT			
Treatments	2	278.6	<0.0001
Residuals	12		
2 WAT			
Treatments	2	216.8	<0.0001
Residuals	12		
3 WAT			
Treatments	2	73.0	<0.0001
Residuals	12		

Table 4.20: Percent weed control of three application treatments, year one.

	Broad		SS		UTC	
	M	*SD*	*M*	*SD*	*M*	*SD*
	--%--					
1 WAT	98.84$_A$	1.69	89.49$_A$	9.41	1.49$_B$	1.69
2 WAT	96.48 $_A$	2.60	88.14 $_A$	10.22	1.49 $_B$	1.69
3 WAT	90.66$_A$	5.95	77.04 $_A$	15.49	1.49$_B$	1.69

Note. Numbers within rows that share subscript letters are not significantly different ($p \geq .05$).
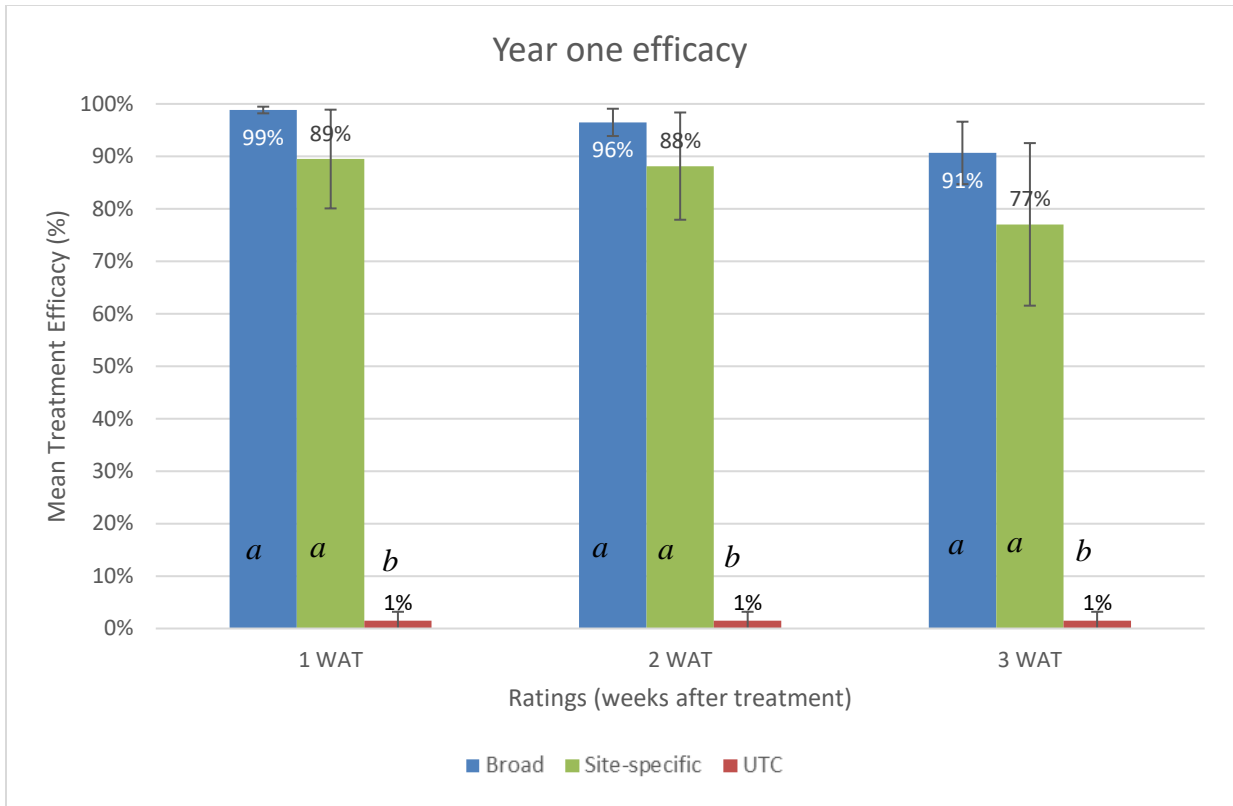
Figure 4.26: Year one treatment efficacy by weeks. Shared letters within weekly ratings indicate a lack of significant (p < .05) difference.

In year one, there was a significant ($p < .05$) increase in weed control (%) between untreated check plots and broadcast plots within each week (97.35%, 94.98%, and 89.17% increase over untreated check one, two, three WAT, respectively). There was also a significant ($p < .05$) increase in weed control (%) between untreated check plots and site-specific plots within each week (88.00%, 86.65%, and 75.54%, increase over untreated check one, two, three WAT, respectively. Broadcast plot means consistently ranked higher than site-specific means (9.35%, 8.34%, and 13.62% for one, two, and three WAT); however, none of these differences proved to be significant ($p < .05$).

Table 4.20 and Figure 4.26 shows that the standard deviations for site-specific plots were much higher compared to those of Broadcast and Untreated check plots. Large variations in data values due to wider variations in weed control, as represented by their standard deviation, require larger differences between means for values to be considered significant ($p < .05$). Therefore, some of the means across within weeks were not different enough to be significant ($p < .05$) with standard deviation values this large.

Treatment Efficacy Year Two

Individual weed control ratings at one, two and three weeks after treatment for year two was averaged by plot. Year two mean weed control (%) was analyzed using an ANOVA. Table 4.21. shows that ANOVA results for weed control were all significant between treatments ($p < .05$) within year two. Once ANOVA results were determined to be significant, the means were then subjected to Tukey's post hoc analysis to determine the amount of significance between each treatment. Means, standard deviations, and significant differences are for year two weed control (%) between treatments are summarized in Table 4.22 and Figure 4.27.

Table 4.21: ANOVA results for year two treatment efficacy.

Year two	*df*	F	*p*
1 WAT			
Treatments	2	291.2	<0.0001
Residuals	18		
2 WAT			
Treatments	2	599.0	<0.0001
Residuals	18		
3 WAT			
Treatments	2	130.5	<0.0001
Residuals	18		

Table 4.22: Percent weed control of three application treatments for year two.

	Broad		SS		UTC	
	M	*SD*	*M*	*SD*	*M*	*SD*
	--%--					
1 WAT	94.53$_A$	3.94	90.71$_A$	9.95	3.92$_B$	6.99
2 WAT	94.16$_A$	6.47	91.07$_A$	6.35	1.16$_B$	2.66
3 WAT	86.97$_A$	13.35	86.53$_A$	9.16	4.25$_B$	9.32

Note. Numbers within rows that share subscript letters are not significantly different ($p \geq .05$).
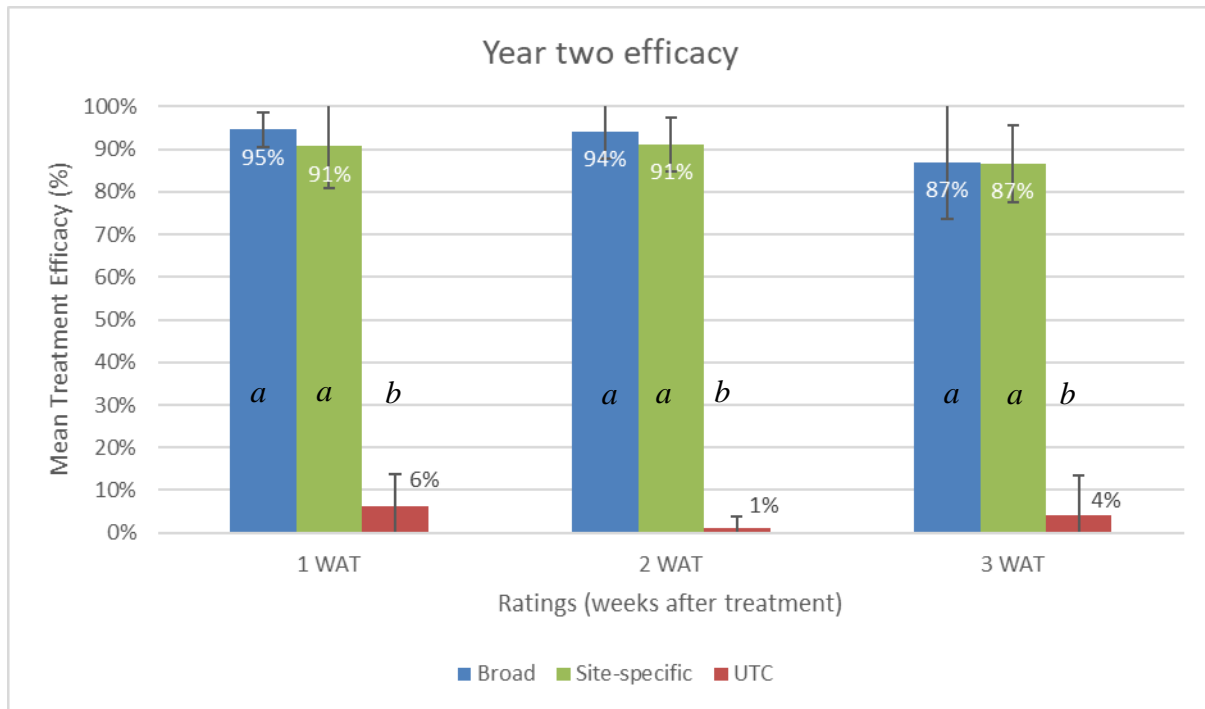
Figure 4.27: Year two treatment efficacy by weeks. Means sharing the same letters within weekly ratings are not significantly (p < .05) different.

In year two, there was a significant ($p < .05$) increase in weed control (%) between untreated check plots and broadcast plots within each week (88.16%, 93.00%, and 82.72% increase over untreated check one, two, and three WAT, respectively). There was also a significant ($p < .05$) increase in weed control (%) in site-specific plots compared to that of untreated check plots both within weeks (84.35%, 89.91%, and 82.28% increase over untreated check one, two, and three WAT, respectively.

Table 4.22 and Figure 4.27 shows that there is a consistent numerical drop in site-specific mean weed control values compared to those of Broadcast (3.82%, 3.09%, and 0.44% for weeks one, two, and three); however, these differences were not statistically significant ($p < .05$).

Treatment Efficacy Year One and Two

     Individual weed control ratings at one, two and three weeks after treatment for year two

one and two was averaged by treatment. Year one and two mean weed control (%) was analyzed

using an ANOVA. Table 4.23 shows that ANOVA results for weed control were all significant

between treatments ($p < .05$) within one-, two-, and three-week ratings. Once ANOVA results

were determined to be significant, the means were then subjected to Tukey's post hoc analysis to

determine the amount of significance between each treatment. Means, standard deviations, and

significant differences are for year one and two weed control (%) between treatments are

summarized in Table 4.24 and Figure 4.28.

Table 4.23: ANOVA results for treatment efficacies from years one and two.

Weeks after treatment	df	F	p
1 WAT			
Treatments	2	566.8	<0.001
Residuals	33		
2 WAT			
Treatments	2	845.2	<0.001
Residuals	33		
3 WAT			
Treatments	2	205.3	<0.001
Residuals	33		

Table 4.24: Percent weed control of three application treatments, year one and two.

Ratings	Broad _M_	Broad _SD_	SS _M_	SS _SD_	UTC _M_	UTC _SD_
	— % —					
1 WAT	96.22 A	3.90	90.10 A	9.96	4.90 B	6.44
2 WAT	95.15 A	5.34	89.64 A	8.58	1.26 B	2.32
3 WAT	88.55 A	11.07	81.79 A	13.05	3.42 B	7.77

Note. Numbers within rows that share subscript letters are not significantly different ($p \geq .05$).
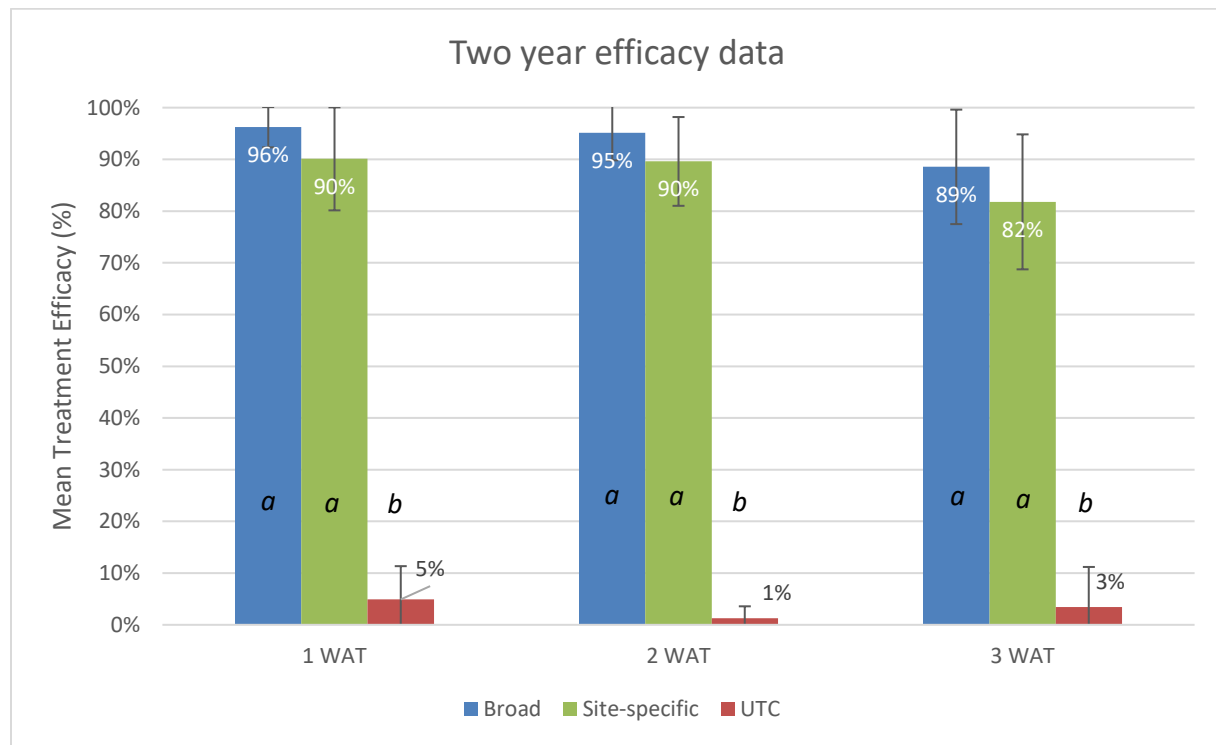

Figure 4.28: Two-year mean treatment efficacy by weeks. Means sharing the same letter within weekly ratings are not significantly ($p < .05$) different.

Like individual years, when years were combined, there was a significant ($p < .05$) increase in weed control (%) between untreated check plots and broadcast plots within each week (91.32%, 93.89%, and 85.13% increase over untreated check one, two, and three WAT,

147

respectively). There was also a significant ($p < .05$) increase in weed control (%) between

untreated check plots and site-specific plots within each week (85.20%, 88.35%, and 78.36%

increase over untreated check one, two, and three WAT, respectively. Despite a consistent

numerical drop in control between broadcast and site-specific plots (6.12%, 5.55%, and 6.77%

for week one, two, and three means), there were no significant ($p < .05$) differences found in

weed control for weekly ratings.

4.5   Discussion

The following section describes the results stated in chapter 4.4, how they relate to

current pesticide application techniques, and their implications for input savings, reductions in

environmental impact. Also discussed are the limitations of these results and the proposed

workflow.

4.5.1   Applied Area

Null hypothesis five stated that there will be no significant ($p < .05$) differences in the

mean treatment area (%) between plots sprayed with a broadcast rate and plots sprayed using

prescription map generated by a SSWM workflow. Significant ($p < .05$) reduction in applied area

were found between broadcast and site-specific treatments (20% and 73% mean reduction in

year 1 and 2, respectively); therefore, null hypothesis five was rejected. Reductions highly

correlated with weed densities within plots. This strongly suggests that if applied at the field

scale it would be possible to significantly reduce the number of hectares treated and the total

volume applied to fields particularly when weed density is low. Water sensitive paper placed

within plots show that areas indicated as untreated areas had no coverage or minimal coverage

from immediately adjacent treaded zones ($M = 3.51\%$, $SD = 6.93$). Water sensitive paper data

also support that those areas designated as treatment areas within site-specific plots (M = 28.95%, SD = 11.30) received statistically similar coverage as the broadcast plots (M = 26.75%, SD = 8.55). This correlation of weed density to sprayed area and spray coverage strongly suggests that the site-specific applications can reduce the environmental impact of pesticides and reduce input cost of producers while maintaining expected coverage of intended targets. Pesticide savings would correlate with weed density and be influenced by weed distribution; however, even with high weed density plots, applied areas were reduced compared to broadcast applications.

4.5.2 Treatment Efficacy

Null hypothesis six stated that plots selectively sprayed using prescription maps generated by a SSWM workflow will not achieve statically ($p < .05$) similar mean weed control compared with broadcast applications. Results showed that there were no significant ($p < .05$) differences in the weekly weed control (%) ratings between broadcast and site-specific herbicide applications within years one and two and across both years of replicated plots. This indicates that the plots selectively sprayed using prescription maps generated by the proposed SSWM workflow achieved statistically similar results as the broadcast treatment leading to the rejection of null hypothesis six. It is worth mentioning a consistent numerical drop in weed control particularly in year one. It is suspected that this difference may be attributed to the statistical ($p < .05$) difference in mean weed size between year one (M = 19.62 cm, SD = 7.19) and year two (M = 13.31 cm, and SD = 4.72) as measured by their detection box size and supported by *in situ* scouting prior to application. This significantly larger set of weeds can be attributed to treatment delays caused by weather in year one and as a group were nearly too large to recommend treating. Herbicide efficacy degrades with an increase in weed size (Meyer & Norsworthy,

149

2019). These degradations may be compounded by slight deviations in coverage or spray volume from turning nozzles on immediately prior to meeting a weed. The greater the distance a weed is from the nozzle, the more time allotted for the spray pattern to fully develop and provide broadcast level coverage. Boom height for each trial was adjusted to the recommended height (0.5 m above target) relative to an approximated mean target height. This was estimated by walking the plots and estimating the weed heights to determining the relative boom height. However, the tallest of weeds (compared with the estimated mean weed height) in year one may have been too close to the nozzles for the broadcast coverage to be achieved. The water sensitive spray cards placed at weed height in year two may have not picked up this difference because of their relative distance from the nozzles since they were placed at weed height. When site-specific treatment maps contain high accuracy detections and are properly geolocated, PWM sprayers can precisely execute the commands and place pesticides accordingly. However, weed size may be a greater concern with site-specific applications than when broadcasting herbicides due to a greater need for nozzles to achieve a full pattern within treatment zones.

4.6   Conclusion

Post-emergent applications of herbicides often apply exceedingly more pesticides than necessary since weeds are spatially sparce. While novel robotic technology has begun to demonstrate capabilities at weed detection and site-specific treatments, today's field application equipment, when equipped with induvial nozzle control via PWM actuators may be able to carry out site-specific application. A site-specific workflow that applied herbicides with a PWM sprayer using prescription weed maps was demonstrated; applied areas and resulting weed control were evaluated.

Significant reductions in applied area were found between site-specific and broadcast treatments. Applied area highly correlated with weed densities within each site-specific plot. The correlation of weed density to applied area and coverage results collected by water sensitive cards, suggests that the site-specific applications can be carried out by current PWM sprayers. Furthermore, site-specific application made by PWM sprayers can reduce the applied area in fields while maintaining expected coverage of intended targets. Herbicide efficacy between site-specific and broadcast treatments were determined to be statistically similar. Assuming site-specific treatment maps contain high accuracy detections that are properly geolocated, PWM sprayers can precisely actuate valves and place pesticides accordingly at the tested speed of 6.4 km/h.

Results of applied area and herbicide efficacy suggest that if PWM sprayers apply prescription weed maps at field scale, similar treatment efficacies may be achieved with significant reductions in the number of hectares treated when compared with broadcast applications. Some variation in nozzle flow rates were noted when engaging different numbers of nozzles, suggesting possible limitations in field equipment's ability to hold target rate across larger sections of plumbing with a single pump. Additionally, PWM sprayers need to be evaluated at their ability to hold uniform applied volumes and resulting coverages at a variety of field speeds.

4.7   References

Akhalaya, B. Kh., & Shogenov, Yu. Kh. (2018). Automated Multifunctional Tillage Machine. *Russian Agricultural Sciences*, *44*(1), 105–107. https://doi.org/10.3103/S1068367418010020

Aravind, K. R., Raja, P., & Pérez-Ruiz, M. (2017). Task-based agricultural mobile robots in arable farming: A review. *Spanish Journal of Agricultural Research*, *15*(1), e02R01. https://doi.org/10.5424/sjar/2017151-9573

Basso, B., Fiorentino, C., Cammarano, D., & Schulthess, U. (2016). Variable rate nitrogen fertilizer response in wheat using remote sensing. *Precision Agriculture*, *17*(2), 168–182. https://doi.org/10.1007/s11119-015-9414-9

Butts, T. R., Samples, C. A., Franca, L. X., Dodds, D. M., Reynolds, D. B., Adams, J. W., Zollinger, R. K., Howatt, K. A., Fritz, B. K., Hoffmann, W. C., Luck, J. D., & Kruger, G. R. (2019). Optimum Droplet Size Using a Pulse-Width Modulation Sprayer for Applications of 2,4-D Choline Plus Glyphosate. *Agronomy Journal*, *111*(3), 1425–1432. https://doi.org/10.2134/agronj2018.07.0463

Campos, J., Gallart, M., Llop, J., Ortega, P., Salcedo, R., & Gil, E. (2020). On-Farm Evaluation of Prescription Map-Based Variable Rate Application of Pesticides in Vineyards. *Agronomy*, *10*(1), 102. https://doi.org/10.3390/agronomy10010102

Castaldi, F., Pelosi, F., Pascucci, S., & Casa, R. (2017). Assessing the potential of images from unmanned aerial vehicles (UAV) to support herbicide patch spraying in maize. *Precision Agriculture*, *18*(1), 76–94. https://doi.org/10.1007/s11119-016-9468-3

Corti, M., Cavalli, D., Cabassi, G., Vigoni, A., Degano, L., & Marino Gallina, P. (2019). Application of a low-cost camera on a UAV to estimate maize nitrogen-related variables. *Precision Agriculture*, *20*(4), 675–696. https://doi.org/10.1007/s11119-018-9609-y

Fernández-Quintanilla, C., Peña, J. M., Andújar, D., Dorado, J., Ribeiro, A., & López-Granados, F. (2018). Is the current state of the art of weed monitoring suitable for site-specific weed management in arable crops? *Weed Research*, *58*(4), 259–272. https://doi.org/10.1111/wre.12307

Gai, J., Tang, L., & Steward, B. L. (2020). Automated crop plant detection based on the fusion of color and depth images for robotic weed control. *Journal of Field Robotics*, *37*(1), 35–52. https://doi.org/10.1002/rob.21897

Gonzalez-de-Soto, M., Emmi, L., Perez-Ruiz, M., Aguera, J., & Gonzalez-de-Santos, P. (2016). Autonomous systems for precise spraying – Evaluation of a robotised patch sprayer. *Biosystems Engineering*, *146*, 165–182. https://doi.org/10.1016/j.biosystemseng.2015.12.018

Kitić, G., Krklješ, D., Panić, M., Petes, C., Birgermajer, S., & Crnojević, V. (2022). Agrobot Lala—An Autonomous Robotic System for Real-Time, In-Field Soil Sampling, and Analysis of Nitrates. *Sensors*, *22*(11), 4207. https://doi.org/10.3390/s22114207

Lehnert, C., McCool, C., Sa, I., & Perez, T. (2020). Performance improvements of a sweet pepper harvesting robot in protected cropping environments. *Journal of Field Robotics*, rob.21973. https://doi.org/10.1002/rob.21973

Luck, J., Sharda, A., Pitla, S., Fulton, J., & Shearer, S. (2011). A Case Study Concerning the Effects of Controller Response and Turning Movements on Application Rate Uniformity with a Self-Propelled Sprayer. *Transactions of the ASABE*, *54*(2), 423–431. https://doi.org/10.13031/2013.36445

Meyer, C. J., & Norsworthy, J. K. (2019). Influence of weed size on herbicide interactions for Enlist[TM] and Roundup Ready® Xtend® technologies. *Weed Technology*, *33*(4), 569–577. https://doi.org/10.1017/wet.2019.27

Oktarina, Y., Dewi, T., Risma, P., & Nawawi, M. (2020). Tomato Harvesting Arm Robot Manipulator; a Pilot Project. *Journal of Physics: Conference Series*, *1500*(1), 012003. https://doi.org/10.1088/1742-6596/1500/1/012003

Roshanianfard, A., & Noguchi, N. (2020). Pumpkin harvesting robotic end-effector. *Computers and Electronics in Agriculture*, *174*, 105503. https://doi.org/10.1016/j.compag.2020.105503

Ruigrok, T., van Henten, E., Booij, J., van Boheemen, K., & Kootstra, G. (2020). Application-Specific Evaluation of a Weed-Detection Algorithm for Plant-Specific Spraying. *Sensors*, *20*(24), 7262. https://doi.org/10.3390/s20247262

Tatsuno, J., Tajima, K., Kato, M., Faculty of Engineering, Kindai University 1 Takaya Umenobe, Higashi-Hiroshima, Hiroshima 739-2116, Japan, Faculty of Regional Environment Science, Tokyo University of Agriculture 1-1-1 Sakuragaoka, Setagaya, Tokyo 156-8502, Japan, & Faculty of Agriculture, Tokyo University of Agriculture 1737 Funako, Atsugi, Kanagawa 243-0034, Japan. (2022). Automatic Transplanting Equipment for Chain Pot Seedlings in Shaft Tillage Cultivation. *Journal of Robotics and Mechatronics*, *34*(1), 10–17. https://doi.org/10.20965/jrm.2022.p0010

Utstumo, T., Urdal, F., Brevik, A., Dørum, J., Netland, J., Overskeid, Ø., Berge, T. W., & Gravdahl, J. T. (2018). Robotic in-row weed control in vegetables. *Computers and Electronics in Agriculture*, *154*, 36–45. https://doi.org/10.1016/j.compag.2018.08.043

Villa-Henriksen, A., Edwards, G. T. C., Green, O., & Sørensen, C. A. G. (2021). Evaluation of Grain Quality-Based Simulated Selective Harvest Performed by an Autonomous Agricultural Robot. *Agronomy*, *11*(9), 1728. https://doi.org/10.3390/agronomy11091728

5.0   SUMMARY AND CONCLUSION

5.1   Summary of the problem

Using high-resolution imagery collected from a low altitude UAS platform may provide the complete field scouting needed for SSWM. Crop/weed identification within this imagery has been proven possible with today's open-source deep learning models, once properly trained; however, *Palmer amaranth* detection in soybean production fields has not been explored to the author's knowledge. PWM application equipment can respond at a relatively fine resolution (1 m); however, are PWM sprayers able to act on detected objects and achieve acceptable weed control if used in a site-specific herbicide? The convergence of these three technologies may meet the foundational requirements for widespread adoption of site-specific pesticide applications as described in Gebbers and Adamchuk (2010). SSWM may be possible with today's equipment but one major limitation is a lack of robust, reliable, and user-friendly workflows (Sishodia et al., 2020). These experiments studied a proposed workflow that combines the capabilities of these technologies and facilitates access to accurate treatment maps. The successful creation of such a workflow would dramatically reduce the volumes of pesticides applied resulting in reduced input cost for producers and reduce environmental impact of production systems. Furthermore, in the case of selective herbicides or application made in genetically tolerant crops, the precision demands of such a workflow would be much lower and may be feasible in current production systems. However, each of the three converging technologies contributions to such a workflow must be evaluated to ensure the integrity of a resulting applications efficacy. These evaluations should include weeds (Palmer amaranth) detection accuracy of different CNN models, the spatial resolution requirements to meet acceptable accuracies, and the application accuracy of equipment action on prescription maps.

154

5.2   Conclusions

Weed control via herbicide applications in modern production systems have been relatively robust and successful. However, post-emergent applications of herbicides often apply exceedingly more product than necessary because weed are spatially sparce compared. A site-specific application to weeds would strategically target the placement of herbicides with the goal of achieving the same efficacy while treating less area. Such a workflow would need to accurately detect and be able to apply herbicides to weed locations. Experiments to evaluate the detection accuracy, herbicide efficacy, and the difference in applied area between broadcast applications and a novel site-specific herbicide application workflow were conducted.

Sixteen CNN models (RetinaNet, Faster R-CNN, YOLO v3, and SSD each trained on imagery 0.4, 0.6, 0.8, and 1.2 cm spatial resolution) were trained to detect *Palmer amaranth* in soybean field imagery. Precision, recall, F1-score, and detection rate results from each model were calculated for detections made on 7,800 weeds from four production fields. Both CNN architecture and training imagery resolution influenced model results. Detection accuracies increased proportionally, and detection rates decreased proportionally with increasing spatial resolution training imagery. Faster R-CNN across the four spatial resolutions performed the most accurate detections followed by YOLO v3, SSD, and RetinaNet. Of the sixteen models trained, only the combination of Faster R-CNN trained on 0.4 cm GSD proved to meet field ready expectations for precision, recall and their harmonized mean, F1-score. While Faster R-CNN maintained impressive detection metrics (precision = 0.80, recall = 0.98, F1-scores = 0.94) across deployments in four trials over two years, further testing of its ability to generalize is recommended. Specifically, an increase in false positive detections were noted at the edges of fields where soybeans were stunted and smaller. This reduced the precision of some fields and

155

suggested that accuracies in younger soybean fields may need to be assessed. Additionally, it was found that weed heights smaller than 2.5 cm decreased in detection recall. This limitation suggests that workflows using 0.4 cm spatial resolution would be most accurate if targeting weeds ≥ 2.5 cm heights. This threshold describes weeds that escape control in an earlier post emergent application. By targeting only these larger escapes, a site-specific workflow would achieve higher detection accuracies. Detection rates were determined to be marginally practical on coarser imagery; however, high resolution imagery detection times became excessive in the context of a single day turnaround timeline.

Significant reductions in applied area were found between site-specific and broadcast treatments. Applied area highly correlated with weed densities within each site-specific plot. The correlation of weed density to applied area and supplemental spray coverage results collected by water sensitive cards, suggests that the site-specific applications can be carried out by current PWM sprayers. Furthermore, site-specific application made by PWM sprayers can reduce the applied area in fields while maintaining expected coverage of intended targets. Herbicide efficacy between site-specific and broadcast treatments were determined to be statistically similar. Assuming site-specific treatment maps contain high accuracy detections that are properly geolocated, PWM sprayers can precisely actuate valves and place pesticides accordingly.

Results of applied area and herbicide efficacy suggest that if PWM sprayers apply prescription weed maps at field scale, similar treatment efficacies may be achieved with significant reductions in the number of hectares treated when compared with broadcast applications. Some variation in nozzle flow rates were noted when engaging different numbers of nozzles suggesting possible limitations in field equipment's ability to hold target rate across larger sections of plumbing with a single pump. Additionally, PWM sprayers need to be

evaluated at their ability to hold uniform applied volumes and resulting coverages at a variety of field speeds.

5.3   Future work

This research demonstrated a successful workflow consisting of three converging technologies; however, each of these technologies have their own limitations that likely inhibit current widespread adoption, offering opportunities for future development and evaluation. While it was the goal of this research to produce a highly practical and accessible site-specific workflow, it is equally important to highlight the current limitations in need of further development.

It was determined and widely supported in the literature that higher spatial resolution imagery produces higher accuracy detection results. Obtaining these spatial resolution demands limit flight altitudes and speeds. In the case of this research, 0.4 cm spatial resolution imagery was successfully demonstrated; however, this limited flight altitude to 18.3 m and flight velocity to 1.83 m/s with the UAS used. At the necessary 70% side by 80 frontal overlaps required for orthorectifying imagery, a 16-ha field would take approximately four and a half hours to collect imagery. The same field at coarser spatial resolutions become much more practical with 0.6, 0.8 and 1.2 cm spatial resolutions taking 2.0, 1.2, and 0.5 hours, respectively. While this may be possible it is not practical nor easily scalable across large production systems with the equipment used in this experiment. Furthermore, weeds are ideally treated when they are small; however, even at 0.4 cm GSD emerging plants are difficult or impossible to detect. Increasing flight altitude to increase data collection efficiency would coarsen spatial resolution and further reduce the accuracy of small weed detection. Development and evaluation of UAS with higher spatial

157

resolution sensors, advanced collection techniques, and extended flight capabilities may address these issues.

Faster R-CNN maintained impressive detection metrics (80% precision, 98% recall, 94% F1-scores) across deployments in four trials over two years. While these are acceptable levels of detection for field use, further testing of its ability to generalize may need to be explored. Effort has been made to describe the variability contained in each of these field tests so that detection expectations can be set appropriately in any consideration of future testing or deployments. Prior to deployment, models accuracy should be evaluated in fields with soybeans at younger growth stages than that captured by training imagery. Additionally, model accuracy should be assessed on varying soil colors, textures, and moistures. False positive detections were noted at the edges of fields where soybeans were stunted and smaller. This reduced the precision of some fields and suggested that accuracies in younger soybean fields may need to be assessed. This was a limitation of this studies training data and can possibly be broadened with more generalized training samples in future training efforts.

The PWM experimental set up performed as expected at executing site specific maps. It should be noted however, that the trials used to test site-specific maps were relatively small when compared to production systems. These results should be generalized with caution to larger field equipment and fields. It is possible that field computers designed to handle individual polygons as whole field boundaries may struggle when each field instead contains hundreds or thousands of polygon detections. Furthermore, system pressure, herbicide volume and rate may vary more in site-specific application on larger equipment since a single pump is servicing 54 (0.5 m nozzle spacing on 27 m boom) to 72 (0.5 m nozzle spacing on 37 m boom) nozzles.

The acknowledged workflow limitations are very real and likely inhibit current adoption of the workflow in large scale production agriculture. However, each of these technologies are continuing to advance and with this progress a slow diminishment of these limitations is expected.