The Guelph Plot Analyzer: Semi-Automatic Extraction of Small-Plot Research Data from Aerial Imagery

Dylan Drover*, Jacob Nederend[†], Brittany Reiche*, Bill Deen[†], Liz Lee[†] and Graham W. Taylor*[§]

*School of Engineering, University of Guelph, Guelph, Canada
[†]Plant Agriculture, University of Guelph, Canada
[§]Canadian Institute for Advanced Research, Canada
[§]Vector Institute, Toronto, Canada

A paper from the Proceedings of the
14th International Conference on Precision Agriculture
June 24 – June 27, 2018
Montreal, Quebec, Canada

Abstract
Small-plot trials are the foundation of open-field agricultural research because they strike a balance between the control of an artificial environment and the realism of field-scale production. However, the size and scope of this research field is often limited by the ability to collect data, which is limited by access to labour. Remote sensing has long been investigated to allocate labour more efficiently, therefore enabling the rapid collection of data. Imagery collected by unmanned aerial vehicles (UAVs) are a significant development in remote sensing for agricultural research, and their potential for efficient workflows has generated interest from agricultural scientists. However, data analysis techniques have not matured at the same rate, and a knowledge gap exists between end users of the data and those who can manipulate, extract, and deliver it. This study was established to address the barrier to adoption of UAVs created by this knowledge gap. We created a tool that can semi-automatically extract plot-level statistics from UAV-acquired imagery. This tool simplifies tasks that were previously accomplished via a Geographic Information System (GIS) by incorporating these tools into a web-based application, the Guelph Plot Analyzer (GPA). Users can upload a GeoTiff raster file to the application, and are presented with the UAV-acquired map, as well as a variety of polygon drawing tools. Using a hierarchy of Trial to Replication to Plot, the user draws boundaries around each category, and the tool can then automatically populate a shapefile with polygons corresponding to the plots. Polygons can be buffered to remove border effects, and alleyways can be specified to correctly align rows. Once finalized, the user can export the overlay as a shapefile, as well as a spreadsheet containing image statistics, including the mean, median, range, and a histogram of pixel values. The plots are labelled according to the user's specified naming convention, making the data easily transferrable to statistical analysis software, as well as seamless to integrate into existing studies. The plot extraction tool is an efficient means for non-remote sensing scientists to turn qualitative imagery into quantitative measures and will help modernize small-plot research as UAVs become more common.

Keywords
Remote sensing, UAV, small-plot trials

The authors are solely responsible for the content of this paper, which is not a refereed publication. Citation of this work should state that it is from the Proceedings of the 14th International Conference on Precision Agriculture. EXAMPLE: Lastname, A. B. & Coauthor, C. D. (2018). Title of paper. In

Introduction

While commercial agriculture remains the focus of precision agriculture (PA) methods, various tools are of equal value to improving and increasing the outputs of agricultural field research (Shi et al. 2016). Agricultural scientists use small-plot trials to test predictions of crop responses to genetic, agronomic, and physiological conditions in a representative environment. The ability to artificially replicate a field environment while controlling much of the underlying variability that could mask treatment effects makes small-plot research a fundamental tool. There is a compromise, however, between the size and scope of research projects and a researcher's capacity to collect data while maintaining the experimental integrity of the trial and its associated observations. Many observations, such as those in phenotyping, are labour-limited and necessitate either more assistants or smaller experiments (White et al. 2012). Remote sensing has long been investigated to allocate labour more efficiently, therefore enabling the rapid collection of data (White et al. 2012; Shi et al. 2016). The employment of unmanned aerial vehicles (UAVs) is a significant development in remote sensing for agricultural research, and their efficient workflow appears to be suitable for rapid data collection in multiple research applications.

Although UAVs seem ideal for applications in PA, there is a relatively slow uptake of the technology in the industry (Zhang and Kovacs 2012; Freeman and Freeland 2015). The lack of availability of reliable estimates of return on investment from the analysis of remotely sensed images is a primary reason for their low adoption rate (Lambert et al. 2004). Additionally, several limitations of this technology include: the collection and delivery of images in a timely manner, lack of image processing and interpretation software, and the integration of remotely-sensed data with agronomic data into expert systems (Du et al. 2008). There has yet to be a widely-accepted software tool developed for UAVs in small-plot research, and a knowledge gap exists between end users of the data and those who can manipulate, extract, and deliver it. The objective of this research was to develop a tool for plot-level data extraction that agricultural researchers can use independently from the labour of geospatial and programming experts.

To this end, we have developed the Guelph Plot Analyzer (GPA) for the semi-automatic extraction of small-plot research data from aerial imagery. The web-based application allows users to upload UAV-acquired data, and uses a semi-automated approach to segment the image into objects with the hierarchy of Trial to Replication to Plot. The application is compatible with raster images collected using various sensors (e.g. RGB, spectral indices, thermal, LiDAR), and with raster-interpolated point data, such as soil and yield maps. The GPA is capable of extracting data from tens to thousands of plots rapidly and accurately based on the GPS coordinates of the trial, and can label plot-level results according to common labelling conventions. The user can then generate a shape file for future use, and a Comma Separated Values (CSV) file containing some statistical information from each plot. This work also included focus group testing of this tool by several research groups from both industry and academia, which resulted in a unanimous conclusion that all participants would be eager to incorporate this product into their current research workflows.

The employment of this tool is highly-efficient over the current state of the art, which requires graduate students or other personnel to walk the fields themselves and record their results using subjective qualitative measurements. The addition of this application to current research methods improves the speed at which these studies can be conducted, as well as the accuracy and objectivity of results. This tool enables the automation of large crop studies, as further developments that leverage computer vision techniques could include the calculation of several properties, including soil quality, plant coverage and emergence, and plant maturity.

The Guelph Plot Analyzer

The GPA is hosted by the University of Guelph[1,2]. Once a user has logged into the application, they are presented with the view seen in Figure 1. The large blank area is the “canvas” where images are loaded and segmented. Uploaded image names are visible near the bottom-left hand side of the window, under the “Gallery” heading. Clicking on an image name in the Gallery loads the image onto the Canvas, as seen in

[1] https://www.uoguelph.ca/engineering/GPA
[2] Video tutorials for the application are available at https://www.youtube.com/watch?v=4turPrgZoII

Figure 2.

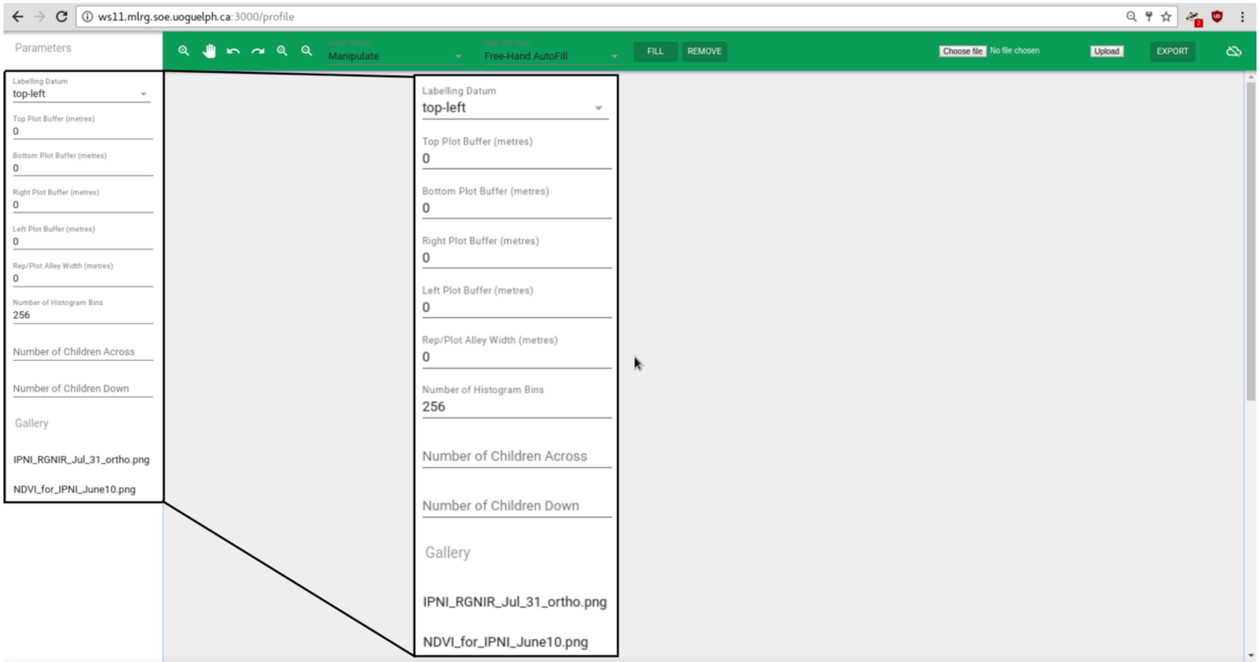


Fig. 1 – View of the University of the Guelph Plot Analyzer after login.

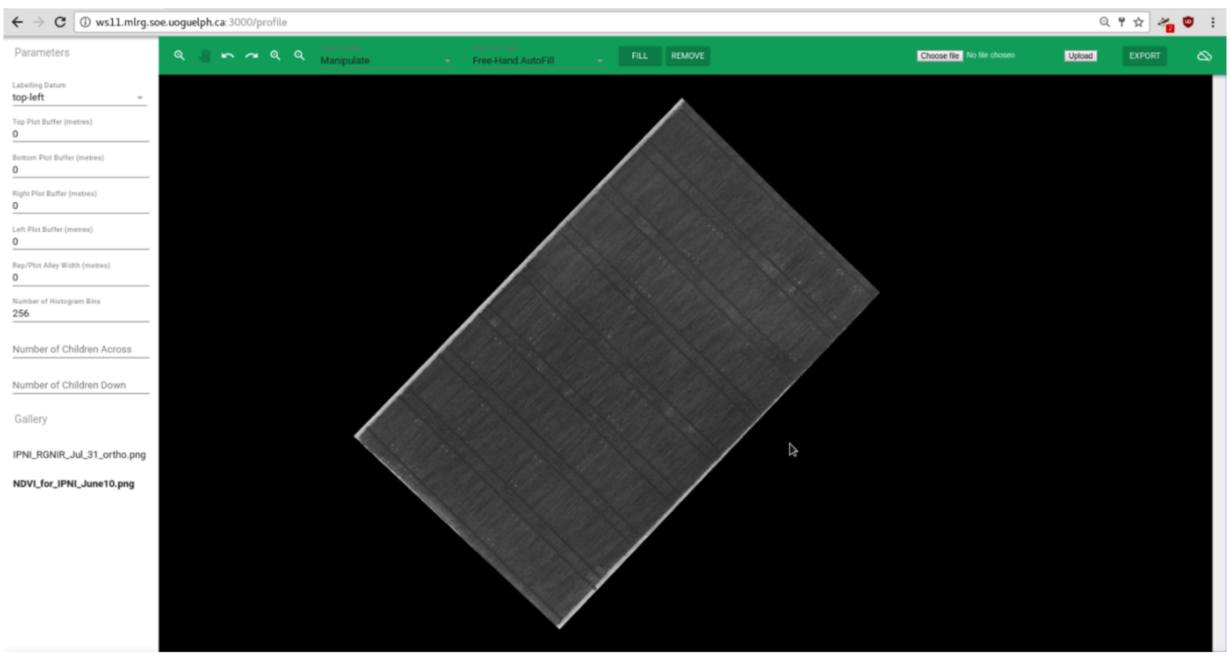


Fig. 2 – View of Selected image loaded from the "Gallery" (visible at the bottom-left hand side of the application).

Initializing the Trial

The first step of segmentation requires the precise positioning of a Trial. Inserting and positioning an "Anchor" creates a snap point and rotation fulcrum for object (Trials, Replicates (Reps), and Plots) corners that have been snapped to it. An example of an Anchor and a Trial are shown in Figure 3. The Trial's top-left corner is snapped to the Anchor and can rotate around that point.

Auto-Fill for Reps and Plots

Once a Trial is positioned correctly, it is populated by Reps through the auto-fill functionality. Figure 4 shows four Reps that populated the initialized Trial, labelled as *100, 200, 300,* and *400*. There are two methods for performing auto-fill: *Free-Hand* and *Set-Number*. These methods are used for creating both the Rep and Plot objects, where Reps are created inside of Trials, and Plots are created inside of Reps.

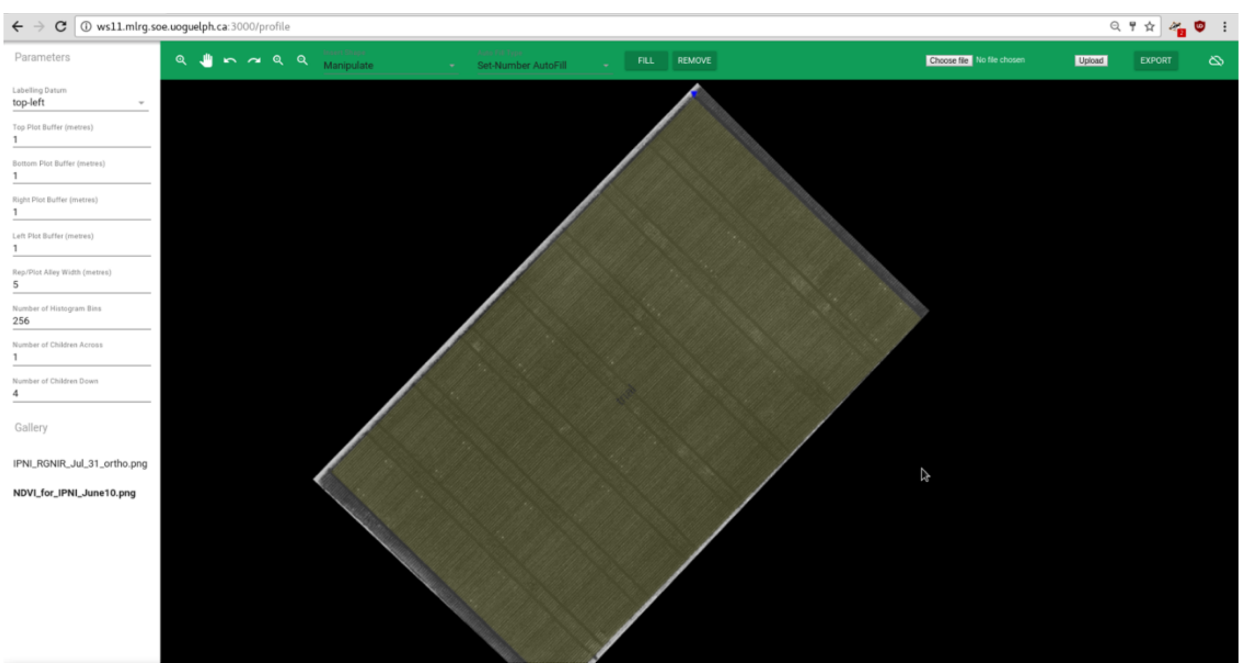
In *Free-Hand* auto-fill, the user can manually draw a Rep within a Trial, and the auto-fill functionality will automatically populate the rest of the Reps in the Trial based on the size of the Rep that was initially drawn, as well as the specified alley width (located in the Parameters side-bar).

In *Set-Number* auto-fill, the user can specify the number of "Children" in which to populate an object, as highlighted in Figure 4. For example, when drawing Reps, the Reps are the Children of the Trial (Parent), and similarly, Plots are the Children of Reps. To create these objects, the user selects the number of Children that go across and down the dimensions of the Parent object, as well as the alley widths between these objects.

Once the Reps have been drawn, the same procedure can be used to create Plots. However, when using auto-fill to populate Plot objects, a buffer parameter can also be specified in the Parameter side-bar, as shown in Figure 5. By specifying this buffer, the central area of each plot is isolated, and this region-of-interest that is used for data extraction (area shown as green in Figure 5). The information located around the edges of each Plot is excluded (shown as white), reducing the noise in the extracted data.

In both cases of auto-fill, the user can specify the numbering convention for Reps and Plots (first entry in the Parameters side-bar). By allowing the user to specify the numbering convention, the GPA can output results that are labelled in a way that is similar to a researcher's existing experiment design. Furthermore, users are not constrained to a rectangle-based case, as has been used as an example here. Additional Trials, Reps, and Plots can be drawn and customized, depending on the shape of the area of study. These objects can also have their numerical-based labels edited individually, making the labelled results compatible with nearly any study.

Fig. 3 – Anchor Snap-point (inverted blue triangle) and positioned Trial.



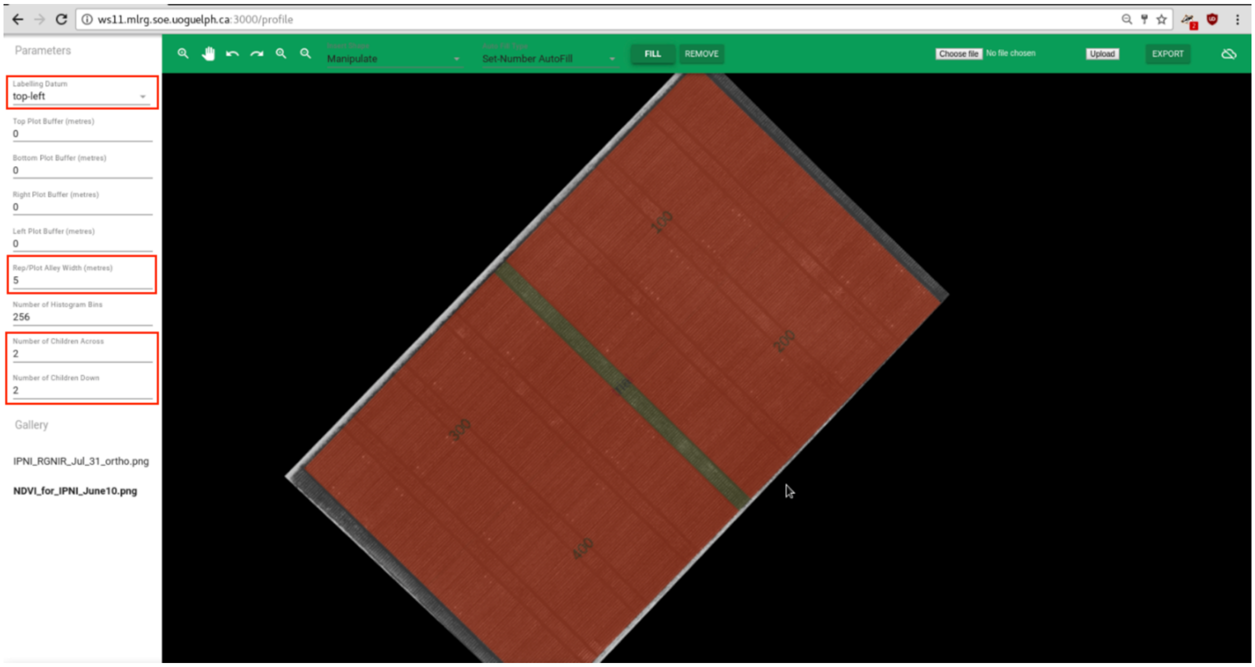


Fig. 4 – View after Rep auto-fill was performed for Trial. Set-Number auto-fill parameters are highlighted in the user menu.

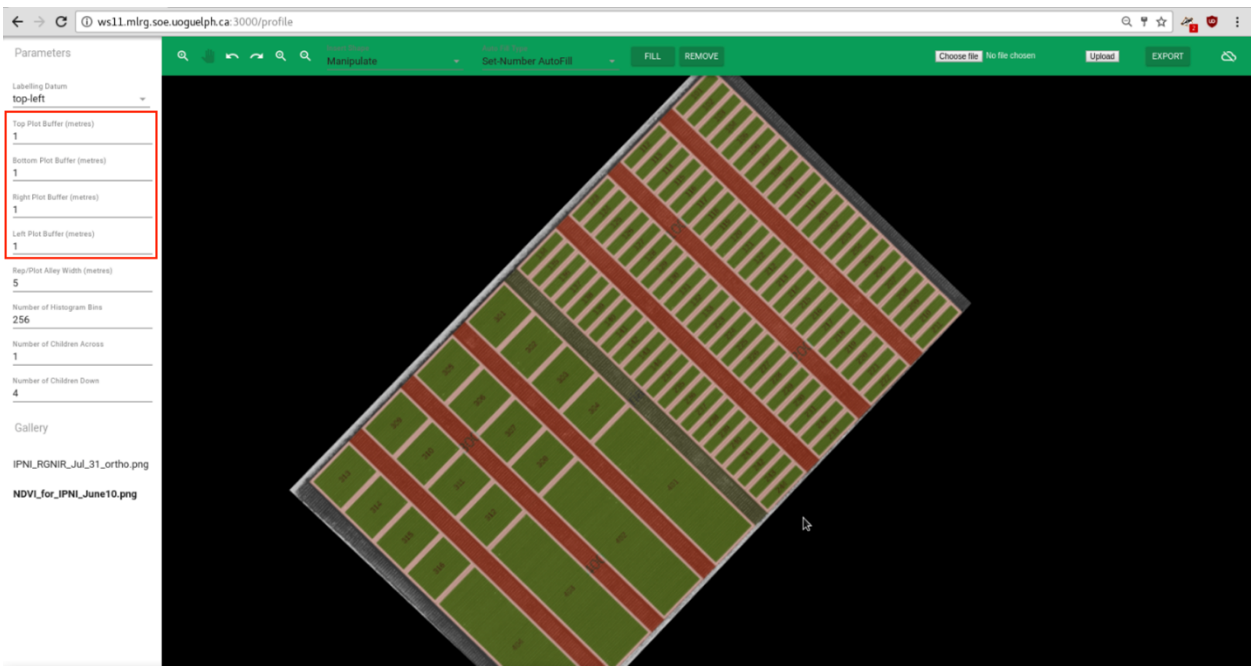


Fig. 5 – View after Plot auto-fill was performed. Each Rep has labelled plots. Plots have 1m buffers (white area) surrounding the area of interest (green).

Exporting Data

When segmentation is complete, the user can select the "Export" button, located in the top-right corner of the application. This immediately downloads a shapefile, which is compatible for use in Geographic Information System (GIS) software packages (for example, ArcGIS[3]). Figure 6 shows a typical visualization of a shapefile. A loading screen will appear and remain until the statistics CSV file has been prepared. The CSV contains plot-level data statistics, including mean, minimum, maximum, and a histogram of the intensity distribution of each plot. This information can easily be imported and analyzed by existing statistical software (e.g. SPSS), or imported by general-purpose programming languages (e.g. Matlab, Python).

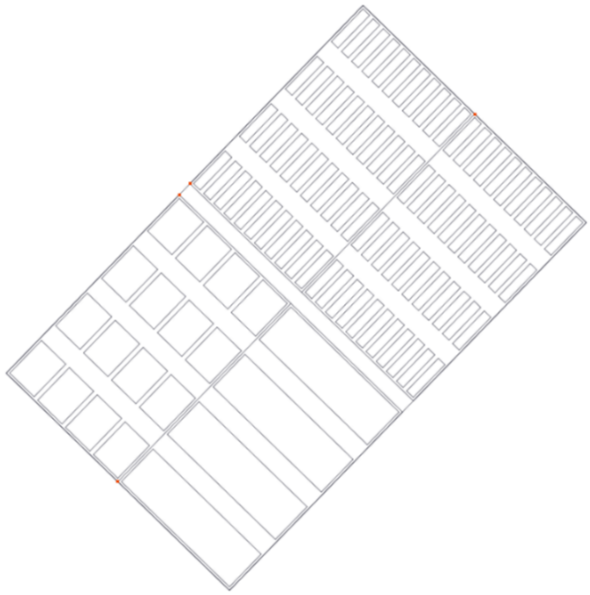


Fig. 6 – Preview of polygons from the shapefile of the segmented Trial from Fig. 4 and 5.

Focus Group Evaluation

A focus group-based evaluation of the application was conducted with research groups from both industry and academia, via remote and on-site studies, to accurately assess the usability of the program, as well as the effectiveness of the video tutorials. Through both venues of evaluation, the participants noted no complaints regarding the usability of the program. In fact, most users were enthusiastic about how quickly and accurately the program could extract plot data, and expressed the desire to apply this to their own research workflows. Participants commented on how the employment of this tool could decrease the length of time required to conduct a study, as well as increasing the objectivity of measurements. The potential reduction of visual ratings, which require strict protocols to reduce bias and require significant labour to complete, was the primary motivator among participants.

Additionally, participants suggested a variety of options for more sophisticated data analysis, including analysis of colour content to determine crop maturity, plant coverage, and emergence, as well as the study of spectrometry data from remote sensing.

Discussion

The GPA is a step towards closing the knowledge gap that restricts the use of UAV imagery in high-volume research workflows. The main priority of this work was the extraction of plot-level data from imagery such that it could be readily compatible with statistical software. Like Shi et al. (2016), an interdisciplinary team of agricultural researchers, computer engineers, and geospatial experts collaborated to ensure a mutual understanding of the needs of the end user and the capabilities of current technology. The objective of the

[3] https://www.arcgis.com

GPA is to enable researchers to extract data without relying on the labour and advice of geospatial and computer programming experts. Any intermediate handling of imagery prior to delivery of the data eliminates the benefits of UAV imagery to researchers by replacing existing labour with that of a third-party. Because of the diverse interests and backgrounds of the collaborating team, the potential for miscommunication and delay of data delivery during transition is high. By collaborating from the beginning, and defining the end users and their needs, we eliminated this source of error by guaranteeing that the tool conforms to the conventions of both geospatial and agricultural work.

The GPA was successful in achieving this objective, based on the results of focus group evaluations. Groups involved included agricultural researchers with familiarities with remote sensing and GIS ranging from none to cursory, in order to demonstrate ease-of-use. Rapid turnaround of plot-level data was of critical importance, and groups unanimously concluded that the GPA could be seamlessly incorporated into their projects. Groups noted that the GPA has limited functionality, but that it addresses the first step in data extraction, which is the digitization of plots in a simple and efficient manner. The use of more sophisticated analytics involving computer vision still require expertise in the realm of geospatial data analysis, but such functionality could be easily integrated with the GPA, owing to its modular and extensible design. These additions must maintain the simplicity of user interactions with the interface and be automated to the greatest extent possible, making minimal expectations with respect to the end user's technical expertise.

Developing robust and accurate methods for image analysis is important for exploiting the image data that can be rapidly collected by UAVs. Some methods already exist, and have been applied to the tasks of detecting stand loss and maize defoliation (Erickson et al. 2004), detecting crop biomass and nitrogen status (Hunt et al. 2005) and predicting soil organic carbon (Gomez, Viscarra Rossel, and McBratney 2008). Shi et al. (2016) integrated UAVs into a high-throughput phenotyping workflow through an interdisciplinary network of experts. Most of these applications require automation of specialized image processing methods so that results can be computed in a reasonable amount of time (Hardin and Jensen 2011). Ideally, the GPA will eventually contain the functionality to compute these same measurements within the same interface.

To improve the functionality of the GPA, future development requires integration of computer vision algorithms, and the ability to import a diverse range of file types. While basic statistics of raster images are useful for correlations to other whole-plot averages, such as yield, more complex measures at the sub-plot level, such as weed coverage or plant count, require more sophisticated processing prior to data extraction. These analytics could be built into the backend of the GPA to avoid exposing the end users to technicalities that fall outside their area of expertise.

Conclusion

This project yielded an accurate, efficient, and robust plot extraction application for remote sensing of small-plot research trials. Focus group testing demonstrated that the program was straightforward and easy to use for participants across different research projects, both from industry and academia. Additionally, participants were eager to implement this program in their current workflows to decrease the length of time to conduct a study, as well as to increase the objectivity of measurements.

In summary, the GPA is a cloud-based software that can automatically extract plot data from aerial imagery, as this service does not yet exist. The program was developed such that it is:

- Web-based;
- Compatible with shapefiles, aerial imagery collected using a variety of sensors (e.g. RGB, spectral indices, thermal, LiDAR, etc.), soil, and yield maps;
- Compatible with planting and numbering conventions used in the majority of all agronomic and breeding trials;
- Capable of extracting data from hundreds of plots quickly (in a matter of minutes) and accurately; and
- Able to produce a histogram summarizing the distribution of information in a plot (e.g. mean, variance, quartiles, etc.).

This application appears to have great market potential, considering the extensive number of agricultural field plots in Ontario, as well as the increasing prevalence of remote sensing technologies in PA.

References

Du, Q., Chang, N., Yang, C., & Srilakshmi, K.R. (2008). Combination of multispectral remote sensing, variable rate technology and environmental modeling for citrus pest management. *Journal of Environmental Management*, 86(1): 14–26.

Erickson, B.J., Johannsen, C.J., Vorst, J.J., & Biehl, L.L. (2004). Using remote sensing to assess stand loss and defoliation in maize. *Photogrammetric Engineering & Remote Sensing*, 70(6): 717–22.

Freeman, P.K., & Freeland, R.S. (2015). Agricultural UAVs in the U.S.: Potential, policy, and hype. *Remote Sensing Applications: Society and Environment*, December(2): 35–43.

Gomez, C., Viscarra Rossel, R.A., & McBratney, A.B. (2008). Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. *Geoderma*, 146(3): 403–11.

Hardin, P.J., & Jensen, R.R. (2011). Small-scale unmanned aerial vehicles in environmental remote sensing: Challenges and opportunities. *GIScience & Remote Sensing*, 48(1): 99–111.

Hunt, E.R., Cavigelli, M., Daughtry, C.S.T., Mcmurtrey, J.E., & Walthall, C.L. (2005). Evaluation of digital photography from model aircraft for remote sensing of crop biomass and nitrogen status. *Precision Agriculture*, 6(4): 359–78.

Lambert, D.M., Lowenberg-DeBoer, J., Griffin, T.W., Peone, J., Payne, T., & Daberkow, S.G. (2004). Adoption, profitability, and making better use of precision farming data. *In* Staff Paper. No. 04–06. West Lafayette, IN, USA: Department of Agricultural Economics, Purdue University.

Shi, Y., Thomasson, J.A., Murray, S.C., Pugh, N.A., Rooney, W.L. Shafian, S., et al. (2016). Unmanned aerial vehicles for high-throughput phenotyping and agronomic research. *PLoS One*, 11(7): e0159781.

White, J.W., Andrade-Sanchez, P., Gore, M.A., Bronson, K.F., Coffelt, T.A., Conley, M.M., et al. (2012). Field-based phenomics for plant genetics research. *Field Crop Research*, 133: 101–112.

Zhang, C., & Kovacs, J.M. (2012). The application of small unmanned aerial systems for precision agriculture: A review. *Precision Agriculture*, 13(6): 693–712.