Delineating wetland vegetation species using a UAV-mounted multi-spectral camera and computer-aided classification

Addison DeShambo[1], Ellie Stewart[2], Sophie Ulik[3], Johanna Williams[2]

Geoscience, University of Wisconsin-Whitewater[1]; Geoscience, Lawrence University[2]; Geography, University of Wisconsin-La Crosse[3]

Abstract

Agricultural runoff can be harmful to the environment by increasing soil erosion and flooding, while adding excess nutrients such as phosphorus to downstream waterways. Stormwater detention ponds are a common way to mitigate flood risk and improve water quality in urban areas. This concept has since been adapted to protect susceptible areas downstream of agriculture. Wetland vegetation can be planted around these ponds to uptake nutrients and slow water flow. Vegetation mapping is a frequent practice within land management to monitor health and species distribution. This process could be less labor-intensive if paired with the use of aerial imagery and computer-aided classification. Unmanned Aerial Vehicle (UAV) acquired imagery allows for high temporal and spatial resolutions that can map the environment in high detail and accuracy. Here we present a detailed workflow for species-level mapping from data acquisition in the field through image processing and analysis.

Introduction

Stormwater detention ponds have been used for over 60 years to mitigate urban runoff; this concept is now being applied to treat agricultural runoff as well. Ponds are constructed to detain water from surface flows or drainage tile networks and to capture eroded sediment and phosphorous from nearby agricultural fields. They can help also to protect sensitive water bodies downstream by taking up nutrients that can cause harmful algal blooms and eutrophication.

The margins of the ponds can be seeded with wetland vegetation or allowed to naturally grow from the existing seed bank. Using the existing vegetation stock is low-cost and minimal maintenance, allowing several ponds to be managed at once. This vegetation helps to uptake nutrients and slow water as it flows through the system, and it mimics a natural wetland that supports diverse flora and fauna by providing habitat. However, since these ponds are unplanted, there is no control over the vegetation that is established. It is possible that some of the vegetation may include invasive species such as phragmites and purple loosestrife. Therefore, a way to map the types of vegetation quickly and accurately will be of great benefit to the pond managers, as monitoring the spread of these species over time is critical to developing a comprehensive management plan.

Funding and resources provided by the Wisconsin Space Grant Consortium (SSI22_1.0) and Lawrence University.

The plants that take up nutrients must be periodically harvested, as they will become a source of phosphorus when they decay. Determining the appropriate time for harvest can be made easier with the use of multi-spectral imagery, as the reflectance of plants change as they start to senesce. Multi-spectral imagery can provide land managers with data that is not visible to the human eye such as red-edge and near-infrared wavelengths, which are highly reflected by vegetation and indicate their health. Additionally, aerial imagery allows for the entire site to be surveyed at once, instead of having to traverse the site on foot through dense vegetation or make use of watercraft. Our objective was to develop a workflow that optimizes UAV flight parameters and ArcGIS Pro classification protocols to determine the most accurate process for mapping wetland vegetation using aerial imagery.

Materials and Methods

Study Site. The study site is an agricultural runoff pond and wetland system located in Buchanan, WI (44.268056° N, 88.199267° W). This system was designed by the Outagamie County Land Conservation Department in 2017 to improve water quality and protect downstream waterways such at Plum Creek, a tributary of the Fox River, which flows into Green Bay and Lake Michigan (Figure 1). The pond and wetland cover about 0.27 hectares (0.66 acres), with a drainage area of 8 hectares (20 acres); 6.5 of which are agricultural and the other 1.5 are residential. According to Hess et al. (2003), The dominant source of phosphates in surface water are phosphate-bound soil particles. Although they are not very soluble, studies have shown them to be transported by groundwater. Runoff flows into the lower sediment basin 1, where most of the eroded soil and phosphorus (soil-bound or dissolved) settles out. The water then continues to flow through the wetland cells and northern sediment basin 2 (shown in Figure 1) where the rest is captured. The gravel spreader bars also act to distribute water flow to avoid channel formation throughout the site. The Nature Conservancy received a grant to survey the site from 2018-2020 from which they compiled a comprehensive list of the different flora and fauna observed at the pond over time, demonstrating the rapid growth and establishment of species and the relationships between them. The United States Geological Survey (USGS) has also been monitoring the flow and composition of incoming water (surface and groundwater) as it enters and exits the site, particularly during storms. This monitoring tracks the efficiency of the pond to remove nutrients and suspended sediment from the water.

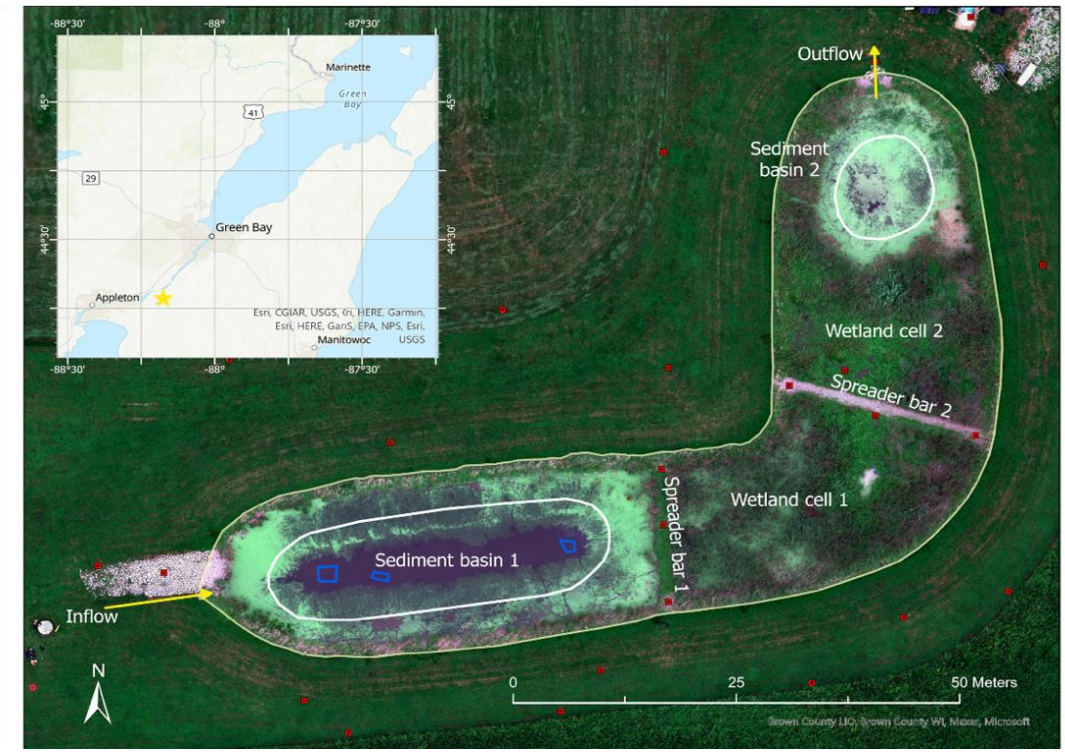


Figure 1. Diagram of the water treatment pond. The yellow star on the inset map indicates the location in Buchanan, Wisconsin. GCPs are represented by red square targets, and the area of interest is outlined with a light-yellow line. Blue polygons in sediment basin 1 show examples of training sites for the class "water".

Field Methods. This study utilized a DJI Matrice 200 V2 drone paired with a multi-spectral MicaSense Altum 8mm camera which takes images in six different wavelength regions (red, green, blue, red-edge, near-infrared, and thermal). There was also a downwelling light sensor that measured ambient light and sun angle to correct for lighting changes that occurred mid-flight. This was paired with the use of a spectral calibration panel before and after each flight, which is used to calibrate the imagery during processing. Imagery was collected on seven different dates between June 22nd and August 1st to assess varying flight altitudes (20m, 30m, and 50m), path orientations (east-west/north-south), and lighting conditions (full sun, overcast, morning/evening) to determine optimal flight parameters. We found the highest quality orthomosaic (ground sampling distance - GSD < 1cm) resulted from a 20m flight altitude with 80% front and 75% side overlap. Having sufficient image overlap is important for computer photogrammetry. It was also found that the even, diffuse lighting conditions afforded by overcast weather were ideal for data collection to avoid shadows and sun glare which are common on clear days. With these flight parameters, there was also a reasonable image processing time of about 12 hours, which means that data can be collected during the workday and processed overnight.

The flight images used for the data analysis in this study were collected on June 30th and August 1st of 2022. In conjunction with the flights, we surveyed 31 ground control points (GCPs – shown in Figure 1) of which 22 were used for georeferencing, and the other 9 were used to validate the model results. The number and placement of the GCPs was determined based on the work of Chandler et al. (2018) that provided a GCP to image ratio and found that even dispersion of GCPs resulted in doubled accuracy in comparison to poor placement. Our GCP distribution was

planned using ArcGIS Field Maps to ensure even and consistent spacing throughout the study site, which also provided a systematic approach to putting them out during each flight. The surveying was done using an Emlid Reach RS2 rover in real-time kinematic (RTK) mode and base established over known coordinates. All points were taken in a "Fix" status, meaning that the rover achieved the solution using all corrections from the base, which yields centimeter accuracy. To establish the base station, it was set to log data for about 5 hours which was an optimal amount of time determined from the literature (Grüner, 2020). The log was then submitted to Natural Resources Canada (NRCAN) for data post-processing and precise point positioning (PPP) in rapid mode for maximum accuracy.

For accurate representation of the vegetation species found at the study site, we met with the county biologist to identify vegetation in the field. Sparse individual plants were left out, as they would be more difficult to see in the imagery, and they would potentially add noise into the classification. Instead, we focused on the most important and dominant species, and created 14 classes (shown in Table 1) that were later combined into 11 overall classes for simplification. Classes that were spectrally similar (e.g., rice cutgrass and pink weed) and classes within the same species that were spectrally different (e.g., dead versus live cattails) were combined (Table 1).

Class	Scientific Names	Training site pixels (%)
Thick Duckweed	*Lemna minor*	10.45
Thin Duckweed		11.71
Narrow-Leaved Cattails	*Typha angustifolia*	20.87
Dead Cattails		0.63
Rice Cutgrass	*Leersia oryzoides*	7.59
Pink Weed	*Persicaria pensylvanica*	5.24
Foxtail barley	*Hordeum Jubatum*	4.45
Purple Loosestrife	*Lythrum salicaria*	0.23
Northern Water Plantain	*Alisma triviale*	2.57
Phragmites	*Phragmites australis*	0.98
Swamp Verbena	*Verbena hastata*	0.22
Soft-Stem Bulrush	*Schoenoplectus validus*	13.34
Rocks		7.09
Water		14.63

Table 1. 14 total classes used for classification. Green, brown, and blue highlights show the 3 groups that were combined. Training site pixels (%) shows the pixel percentages of each class's training sites out of the total amount.

Lab Methods. The image processing and analysis workflow is shown schematically in Figure 2. Agisoft MetaShape was used for radiometric calibration, georectification of the images using GCPs, and creation of orthomosaics. To create an orthom2osaic, the software looks for points of contrast (key-points) that can be found in multiple images, also known as tie-points. These tie points are used to stitch the images together, which is why having sufficient image overlap is important. GCPs are used to reference the photos to real-world coordinates, which allows us to accurately stack and compare orthomosaics from different dates.

The orthomosaic was uploaded into ArcGIS Pro 2.7.2, where training sites were made for each class. These sites are manually created polygons (water class examples shown in Figure 1) that contain spectral information for the computer to recognize each class. The software then classifies the rest of the image (non-training sites) through comparison with the spectral profiles of each training site. Training site coverage for each class can be seen in Table 1. These sites were also used to generate spectral signatures that were used to analyze differences and similarities in reflectance between classes, and to choose the spectral bands that maximize those differences for the best classification results.

Three different classification methods were evaluated to determine which was the most accurate: Deep Learning, Unsupervised, and Supervised. Deep learning was reported to be an effective way to incorporate more than three bands into one classification, which is a notable limitation in the ArcGIS Pro framework (Abd-Elrahman et al., 2021); however, we found the results were poor. Only the dominant species with the largest training site pixel percentage would be classified satisfactorily, with misclassification being common among the classes with a smaller training site pixel percentage (Table 1). Increasing the area of the smaller training sites may improve the result, but we did not test this.

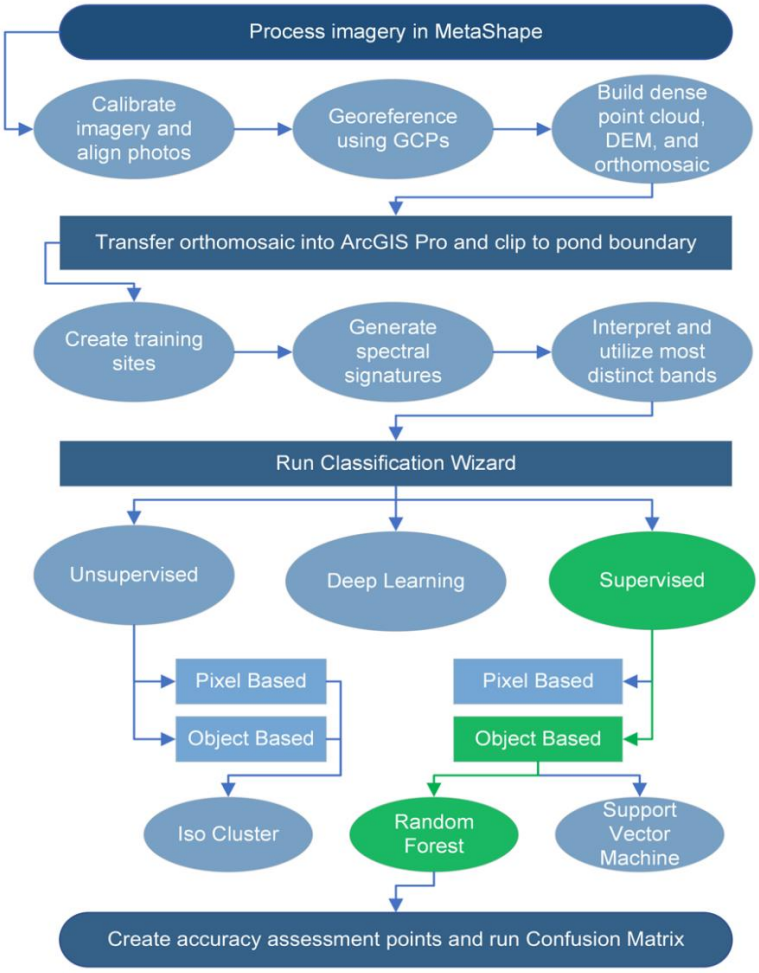


Figure 2. Methodology flow chart. Green shading indicates the path taken though the Classification Wizard in ArcGIS Pro using supervised object-based learning and the random forest algorithm.

ArcGIS Pro has two primary classification schemes: supervised and unsupervised. The unsupervised creates its own image classes by classifying together like objects according to a set number of max potential classes while the supervised uses premade classes and training sites. Within each of these schemes the computer can use a pixel or object-based approach. Object-based classification involves segmentation, where pixels are grouped into clusters or "objects" of similar pixels, and then those objects are classified. Pixel-based approach looks only at the individual pixel and disregards its location and neighbors. The pixel-based approach was not the best fit for this study as it introduced speckle and noise into the classification, also known as the "salt-and-pepper" effect (Blanchard et al., 2011). This is because neighboring pixels of the same vegetation type can be highly contrasting, leading to higher misclassification.

The segmentation parameters were set to have high spectral detail and low spatial detail. This tells the software that pixels must be highly alike to be grouped together into "objects", and that not all similar features are necessarily clustered together.

Unsupervised was not ideal for this study as seasonal changes and varied lighting conditions may cause significant differences in classification results. This makes the results inconsistent and non-repeatable. After comparing numerous iterations, supervised object-based was settled on as this provided the most accurate results. The supervised classification method allowed for application of the random forest algorithm, whereas the unsupervised method was restricted to the Iterative Self Organizing (ISO) scheme. Similar studies achieved superior results using random forest (Cai et al., 2021; Ahmed & Franklin, 2018; Lu & Yuhong, 2017). In this study, the random forest parameters were set to have a high number of trees (300), this means the algorithm ran 300 different decision tree models at once to explore all the possible classification outcomes for a given pixel. Having a higher number of trees means that the algorithm has considered more possibilities, making the classification more informed and accurate. We also increased the tree depth from the default of 30 to 60, this is related to how many times each decision tree is split. Having a low value can result in an over-simplified or biased analysis, while if the value is too high it can make the analysis too complex which adds noise into the classification.

Results

Spectral Signatures. To determine the optimal bands to use for classification, the spectral signatures for the 14 different vegetation classes were analyzed. These spectral signatures show the reflectance of the classes across five band wavelengths for reflected light. The thermal band was not included because of the poor spatial resolution and topographic effects, where vegetation on the south facing slopes warms more than the others. This effect would confuse the classifier. Figure 3 demonstrates the similarities and differences across the bands for five land cover types. The ordinate shows radiometrically calibrated spectral reflectance values, and the abscissa shows the different bands of light the instrument records. Note that all plant types exhibit low reflectance in blue and red wavelengths, while reflecting the green wavelength; this is why plants appear green to our eye. The dolostone rocks used in the inflow area and spreader bars in contrast reflect roughly equally in all visible wavelengths and therefore appear white (Figure 1). Interestingly, plants are very highly reflective beyond the visible range in the red-edge (RE) and near infrared (NIR) wavelengths. It is in these longer wavelength regions that the largest differences in reflectance were observed between each of our classes. Having large spectral differences in the classes is important because it aids the computer

in distinguishing between them. For example, the Phragmites and Rice Cutgrass are similar in the visible range and cannot be distinguished based on color alone, however there is some separation of the reflectance in the NIR and RE regions. Based on these observations we chose to use the NIR and RE bands for our classification input.

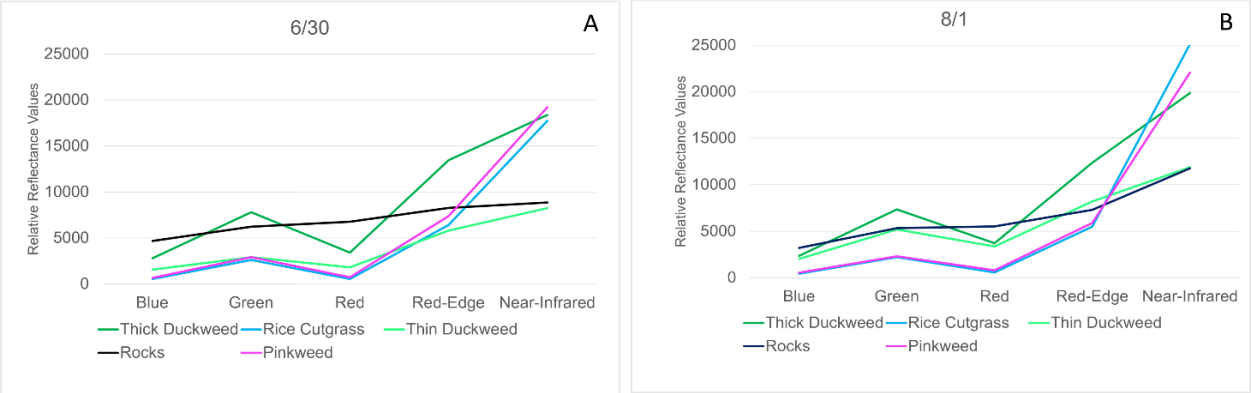
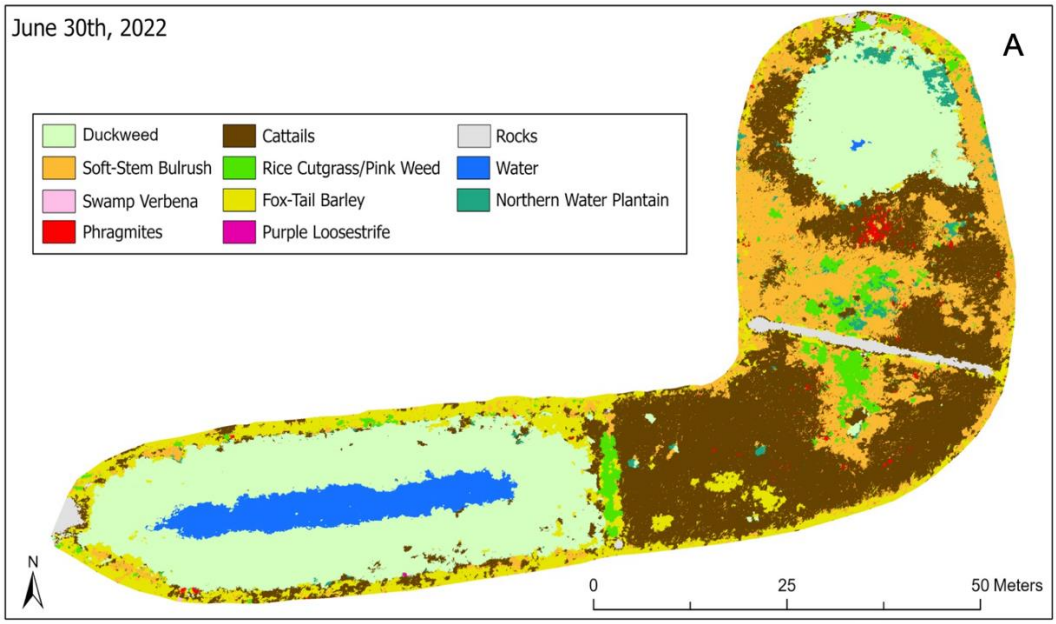


Figure 3. Spectral profiles for five different classes from imagery collected on June 30, 2022 (A) and August 1, 2022 (B).

Classification Maps. The classified orthomosaics show the distribution of wetland plants at each date for which data was collected (Figure 4). By comparing maps on different dates, we can observe changes over time. For the June 30th flight (Figure 4a), there are three dominant species: duckweed, cattail, and bulrush. By August 1st, the bulrush was overtaken by rice cutgrass and pink weed (Figure 4b). Additionally, there was an expansion of duckweed, rice cutgrass and phragmites, the latter of which is a problematic invasive species. Detection of phragmites was an important finding, as the cluster was hidden when observing from the ground due to surrounding vegetation. This suggests that use of aerial imagery can help to reveal hidden species that may have been challenging to reach on foot.



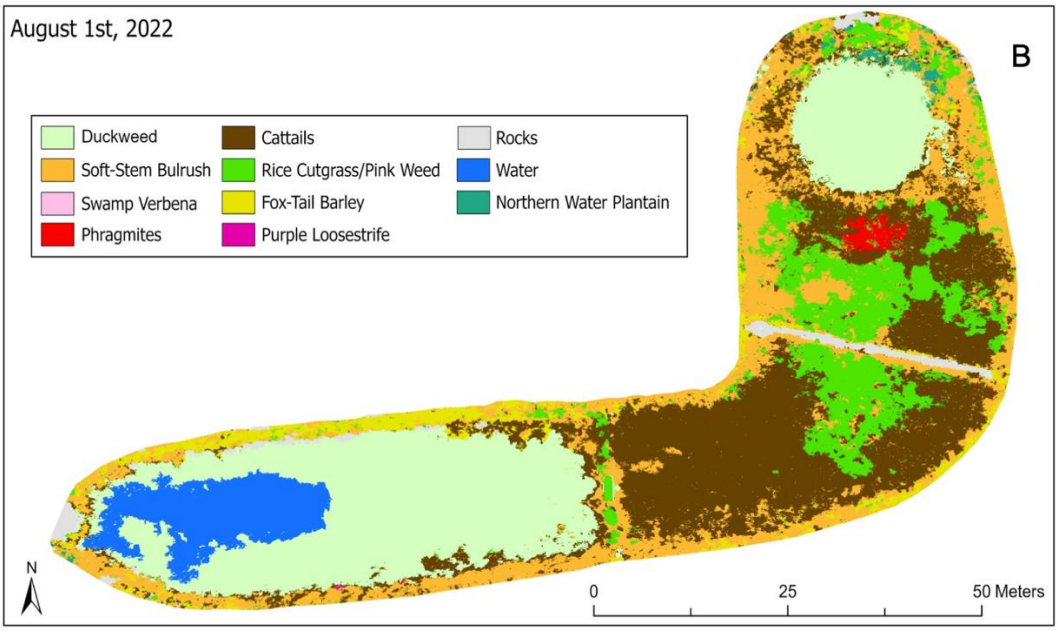


Figure 4. Classification output of June 30th (A) and August 1st (B) imagery using supervised object-based, random forest algorithm, and near-infrared and red-edge bands in ArcGIS Pro.

Accuracy Assessment. To quantify the accuracy of the classification, the classified and actual values for 140 stratified randomly generated points were manually compared. From this, a confusion matrix was created which helps visualize and inform which points were accurately classified (on-diagonal) and which ones were misclassified (off-diagonals) within each class (Table 2). The overall accuracy is a ratio of the number of correctly classified points to total number of points (shown where user's and producer's accuracy meet in Table 2). The kappa reflects the difference between observed accuracy and chance accuracy. For example, a kappa of 0.84 means there is 84% better agreement than by chance alone. Additionally, user's and producer's accuracy give individual accuracy measures for each class. User's (rows) is associated with the real ground cover and error of commission (inclusion); for example, the cattail row in Table 2 shows that 19 pixels were correctly identified as cattails, while the other 3 were incorrectly included in that class when they were truly barley, phragmites, and bulrush. The user's accuracy for cattails then is a ratio of correct pixels to total pixels claimed to be in that class, so in this case, 19/22 = 0.86. On the other hand, producer's accuracy (columns) is associated with the classification schema and error of omission (exclusion), so the cattail column shows that 19 pixels were correctly identified as cattails, while 9 were incorrectly excluded from cattails and assigned to other classes.

After determining high accuracy for the June 30th flight, the same classification was applied to the August 1st flight to assess its reproducibility. We found this workflow to be robust as the imagery from August 1st yielded similar accuracies of 81% total and a kappa of 0.78. The accuracies attained from these classifications are comparable with other literature findings (Cai et al., 2021; He & Lu, 2017) which involved similar analysis and vegetation classification.

Class	Duckweed	Cattails	Cutgrass	Barley	Loosestrife	Rocks	Water	Plantain	Phragmites	Verbena	Bulrush	Total	User's Accuracy	Kappa
Duckweed	26	0	0	0	0	0	0	0	0	0	0	26	1	0
Cattails	0	19	0	1	0	0	0	0	1	0	1	22	0.86	0
Rice Cutgrass	0	0	10	0	0	0	0	0	0	0	0	10	1	0
Fox-Tail Barley	1	2	1	6	0	0	0	0	0	0	0	10	0.6	0
Loosestrife	0	0	0	0	10	0	0	0	0	0	0	10	1	0
Rocks	0	0	0	0	0	10	0	0	0	0	0	10	1	0
Water	0	0	0	0	0	0	10	0	0	0	0	10	1	0
Plantain	0	2	3	0	0	0	0	5	0	0	0	10	0.5	0
Phragmites	0	4	0	1	0	0	0	0	5	0	0	10	0.5	0
Verbena	0	0	1	0	0	0	0	0	0	9	0	10	0.9	0
Bulrush	0	1	1	0	0	0	0	0	0	0	10	12	0.83	0
Total	27	28	16	8	10	10	10	5	6	9	11	140	0	0
Producer's Accuracy	0.96	0.68	0.63	0.75	1	1	1	1	0.83	1	0.91	0	0.86	0
Kappa	0	0	0	0	0	0	0	0	0	0	0	0	0	0.84

Table 2. Confusion matrix for the 06/30 classification with an overall accuracy of 86% and kappa of 0.84.

Discussion

Although we have developed a workflow to map wetland vegetation quickly, consistently, and accurately, two shortcomings to this method were discovered. The large amount of spatial and spectral variation in the plants over time requires manual interpretation and input to maintain training sites for the algorithm. For example, some of the water training site examples shown in Figure 1 worked for classification in late June (Figure 4A), but not in early August (Figure 4B), since the duckweed has expanded substantially over the water. The spectral signatures of land cover also change over time. The thin duckweed and rocks are distinct in the Red-edge and NIR bands in June (Figure 3A) but grew more similar over time (Figure 3B), which makes it difficult for the computer to distinguish them. The spatial and spectral changes over time mean that this workflow cannot be fully automated, however updating the training sites is a relatively straightforward adjustment to make.

Another limitation of this study was the classification scheme within ArcGIS Pro, which can only use 3 spectral bands (rather than the 6 recorded by the instrument). More flexible software could manage additional bands as well as including topographical information; however, we found our results yielded comparable accuracies to studies that did incorporate additional data (e.g., Heilman et al., 2018). Further research could explore these possibilities, but nonetheless, this study shows that the use of simpler software can still yield trustworthy results.

While the main purpose of these ponds is to improve water quality, tracking vegetation at these sites is important for land management and can be used to inform stakeholders on the natural establishment of wetland species, and to monitor the spread of invasive species. The significant spatial and spectral changes over the course of a month suggest that change can be detected and monitored throughout the growing season and likely over a period of years. The ability of geo-rectified UAV-imagery to map these changes supports future application of this technique for wetland vegetation mapping in larger ponds, which may not be as easily accessible by other means.

Conclusion

During this study, our team worked to find optimal flight parameters for our goal of species-level vegetation classification. It was found that a low altitude height (20m) produced a highly detailed orthomosaic with resolution of less than 1 cm per pixel at a cost of about 12 hours of processing time. Overcast weather for flights provided even, consistent lighting and avoided the shadows

and sun glare that occur on a cloudless day. Adding GCPs created an elevated level of geospatial accuracy which allowed for straightforward comparison between different sets (dates) of data. We then found that the red-edge and near-infrared bands provided the highest amount of spectral spread for distinguishing between classes. By combining the determined flight parameters and bands with supervised object-based classification and random forest algorithm, we were able to achieve our highest classification accuracy of 86%. The accuracy of these classifications can vary based on changes in plant growth patterns and spectral variations that occur as the plants mature, however minor adjustments in the training sites can compensate for this. Overall, this process was found to be a robust and reproducible tool for monitoring and delineating wetland vegetation species.

Acknowledgments

We would like to thank the WSGC and Professor Jeff Clark for this opportunity, and Outagamie County Conservation Department for providing background information and access to the study area.

References

Abd-Elrahman, A., Britt, K. & Liu, T. (2021). Deep learning classification of high-resolution drone images using the ArcGIS Pro software. *UF/IFAS Gulf Coast Research and Education Center, 2021*(5), 1-7. https://doi.org/10.32473/edis-fr444-2021

Ahmed, O. S. & Franklin, S. E. (2018). Deciduous tree species classification using object-based analysis and machine learning with unmanned aerial vehicle multispectral data. *International Journal of Remote Sensing, 39*, 5236-5245. https://doi.org/10.1080/01431161.2017.1363442

Blanchard, S. D., Kelly, M., Kersten, E. & Koy, K. (2011). Terrestrial remotely sensed imagery in support of public health: new avenues of research using object-based image analysis. *Remote Sensing 3*(11), 2321-2345. https://doi.org/10.3390/rs3112321

Cai, X., Li, E., Xia, Y., Yang, C., Yang, J. & Zhou, R. (2021). Object-based wetland vegetation classification using multi-feature selection of unoccupied aerial vehicle RGB imagery. *Remote Sensing, 13*(23), 1-21. https://doi.org/10.3390/rs13234910

Chandler, J. H., Ordóñez, C., Rodriguez-Pérez, J. R. & Sanz-Ablanedo, E. (2018). Accuracy of unmanned aerial vehicle (UAV) and SfM photogrammetry survey as a function of the number and location of ground control points used. *Remote Sensing, 10*(10). https://doi.org/10.3390/rs10101606

Grüner, C. (2020, April 14). *Study of deviation of PPP with RS2 with different observation-times*. Emlid Community Forum. https://community.emlid.com/t/study-of-deviation-of-ppp-with-rs2-with-different-observation-times/19185

Heilman, P., McClaran, M. P., McVay, J., Nichols, M, Sankey, T. T. & Swetnam, T. L. (2018). UAV hyperspectral and lidar data and their fusion for arid and semi-arid land vegetation monitoring. *Remote Sensing in Ecology and Conservation, 4*(1), 20-33. https://doi.org/10.1002/rse2.44

Hess, K. M., Kent, D. B., LeBlanc, D. R., McCobb, T. D. & Smith, R. L. (2003). *Phosphorus in a ground-water contaminant plume discharging to ashumet pond, cape cod, massachusetts, 1999*. United States Geological Survey. https://pubs.usgs.gov/wri/wri024306/

Lu, B. & Yuhong, H. (2017). Species classification using unmanned aerial vehicle (UAV)-acquired high spatial resolution imagery in a heterogeneous grassland. *ISPRS Journal of Photogrammetry and Remote Sensing, 128*, 73-85. https://doi.org/10.1016/j.isprsjprs.2017.03.011