Classification of Vegetation Species Along the Lower Salt River Using Unmanned Aircraft Systems and Deep Learning Neural Networks

By: Brandi Kapos, Arnold Kedia, Songmei Liao

2020 MAS-GIS Capstone

Supervisors: Jacob Draper, Justin Eddinger, Christopher Updike
Faculty Contacts: Amy Frazier
Arizona State University
Green Drone AZ

Contents

List of Figures

Chapter 1

1.1 Introduction

The Cactus Fire on April 25, 2017 (Arizona Emergency Information Network, 2017, April 28) affected a total of 818 acres of land on the Lower Salt River (LSR), which was dominated by invasive species including salt cedar (*Tamarix spp.*) (Figure 1a) and giant reed (*Arundo donax*) (Figure 1b). While invasive plants can increase the risk of severe wildland fires, plants such as salt cedar are fire adapted and regrow rapidly, contributing to their ability to outcompete and replace native species (USDA, 2014). Invasive plants also reduce biodiversity, decrease the quantity/quality of natural resources, lead to water shortages, increase flooding, and increase pollution (from chemical removal treatments) (Centre for Agriculture and Bioscience International, n.d.).

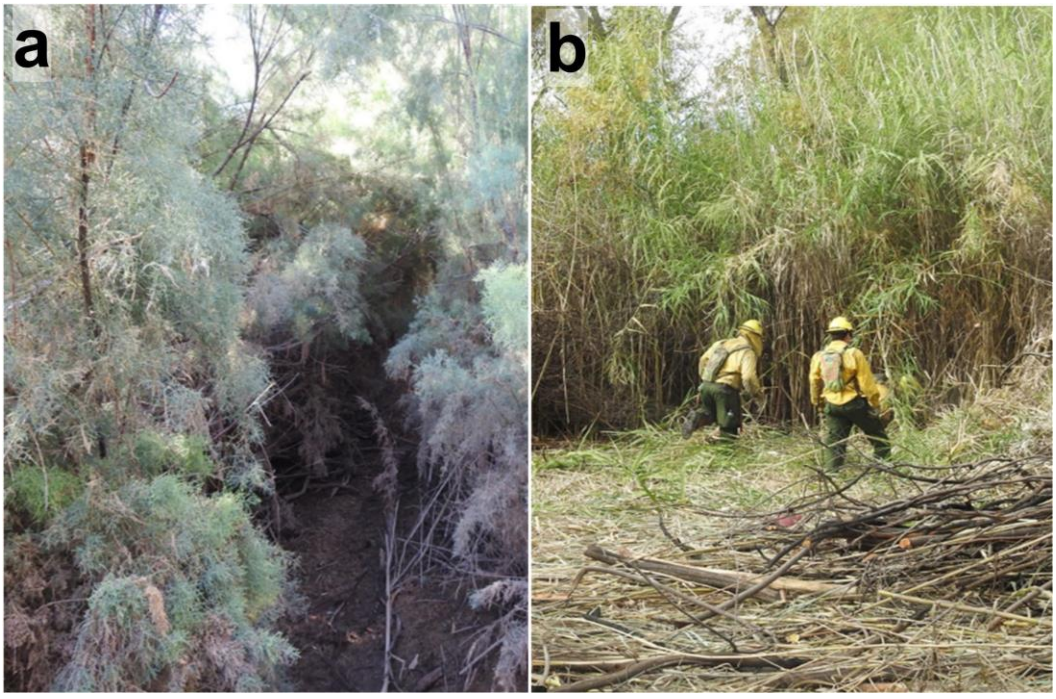

Figure 1. a) Image of the invasive species salt cedar, b) image of restoration workers removing giant reed.

Obtaining funding for restoration projects can be challenging. However, due to the proximity of the LSR to the Phoenix metropolitan area, momentum was quickly gained to develop ecological restoration efforts in the area. These projects have been divided into 46 manageable grids with a mean grid size of 143 acres along the river. So far, approximately 200 acres have ongoing efforts to restore the riparian habitat by removing invasive plants and replanting these acres with 80,000 native trees, including willows (*Salix goodingii* and *S. exigua*), cottonwoods (*Populus fremontii*), palo verde (*Parkinsonia florida*), and ironwood (*Olneya tesota*) trees (National Forest Foundation, n.d.). These replanting efforts are intended to improve wildlife habitat, increase accessibility to the main river channel, and protect the Lower Salt River Recreational Area for all to enjoy.

On flow-regulated rivers, shifts in flood timing favor the reproductively opportunistic salt cedar over cottonwoods and willows, both of which have narrow germination windows. The prevailing hydrological conditions thus favor a new dominant pioneer species in the riparian corridors of the American Southwest (Lite et al., 2007). Flow-regulating dams also influence vegetation by changing soil salinity levels. Due to the altered flow, there have been some unintended consequences for biodiversity, water quality, and ecological processes (Robles et al., 2017). Compared with other ecosystems, rivers support a disproportionately large number of plant and animal species (Ward et al., 1999). The concept of biodiversity and its conservation and restoration is central to maintaining the natural character of rivers and their environment (Addy et al., 2016). To be successful, such efforts must go beyond accounting for flood pulses to restore natural flow variability and achieve hydrological connectivity between a river and its surroundings (Palmer and Ruhi 2019). The quantity and distribution of plant species along the river channel must also be studied and monitored periodically. These data can be used to optimize restoration attempts and help prevent any future fire hazards. Unmanned (Unoccupied) Aircraft Systems (UAS) can be used

to efficiently monitor large areas on a regular basis to stay up to date on restoration efforts (Buters et al., 2019).

1.2 Objectives

The overall objective of this project is to develop a protocol that will help evaluate and monitor the restoration management actions taken along the Lower Salt River using UAS. The specific objectives are:

i. To collect UAS data to create Digital Elevation Models (DEMs) and orthomosaics for hydrology (flood models and surface water feature analysis) and Index layer analysis (Vegetation and soils indices).
ii. To create an inventory of plant species and plant species distribution map using the Random Forest Algorithms.
iii. To create a dashboard that displays the inventory over time.
iv. To create simple protocols for Green Drone AZ to replicate the methods developed to generate these data products and apply them to other areas.

Chapter 2

2.1 Study Location

The study area (Figure 2) is located in the Lower Salt River Recreation Area, which is part of the Tonto National Forest (TNF).

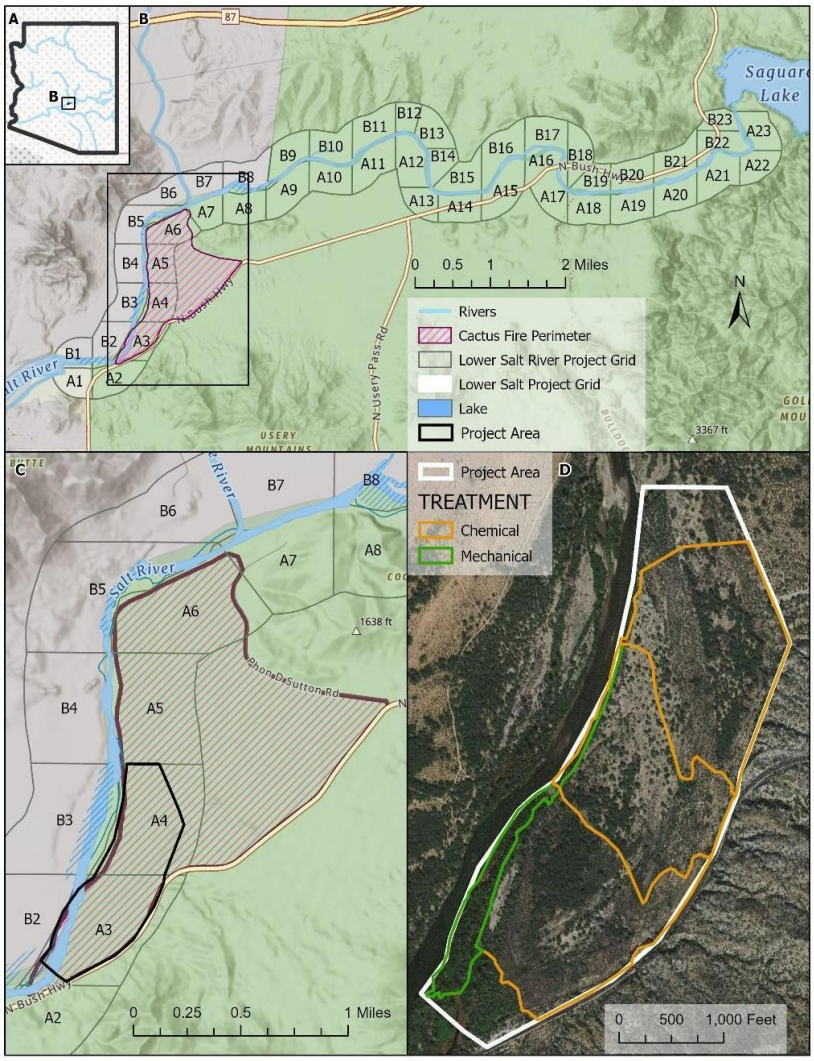

Figure 2. Location of study area. A) A map of Arizona locating the Salt River study area, B) the Lower Salt River Restoration Project divided into 46 grids showing the Cactus Fire perimeter, C) the study area and grid numbers within the cactus fire perimeter, D) aerial imagery (post fire, 2019) for the study area showing the different treatment types applied to the different invasive plant species

This area is approximately four miles northeast of Mesa, Arizona and is one of the most visited "urban" forests in the United States (USDA, n.d.). Climatically, the study area belongs to the desert climate (Köppen climate classification BWh) with 8.4 inches (213.36mm) average annual rainfall ("Mesa, Arizona", n.d., Flood Control District of Maricopa County, 2018). In 2017, the annual rainfall was 9.33 inches (236.98 mm), but there was also no rainfall during the month of April, which means the area was extensively dry (Maricopa County, n.d.). This lack of spring precipitation created an ideal ambient humidity for fire combustion, which contributed to the Cactus Fire. The fire was human induced and quickly spread, causing the National Weather Service to issue a warning for this area after the fire started (Arizona Emergency Information Network, 2017, April 30). According to Mitchell (2017), the humidity was as low as seven percent, and wind speeds were as high as 45 mph (72 kph). Low humidity, high wind speed and clustered dry invasive plants accelerated the Cactus Fire.

The study area for this project covers two of the 46 grids (Figure 2B; A3 and A4) to be restored. The total projected area for this site is 188 acres and includes areas that have already been mechanically and chemically treated. Some parts of this site are difficult to reach due to the dense vegetation, and there is a wildlife protected area to the north. The UAS is an ideal tool to access and collect quality data from such challenging areas where direct access is difficult.

2.2 Methods

Images captured by UAS were used to develop geospatial data products to analyze the physical characteristics of surface water features, flood models, vegetation growth patterns, and vegetation distribution within the study area (Figure 3 and 4)**.**

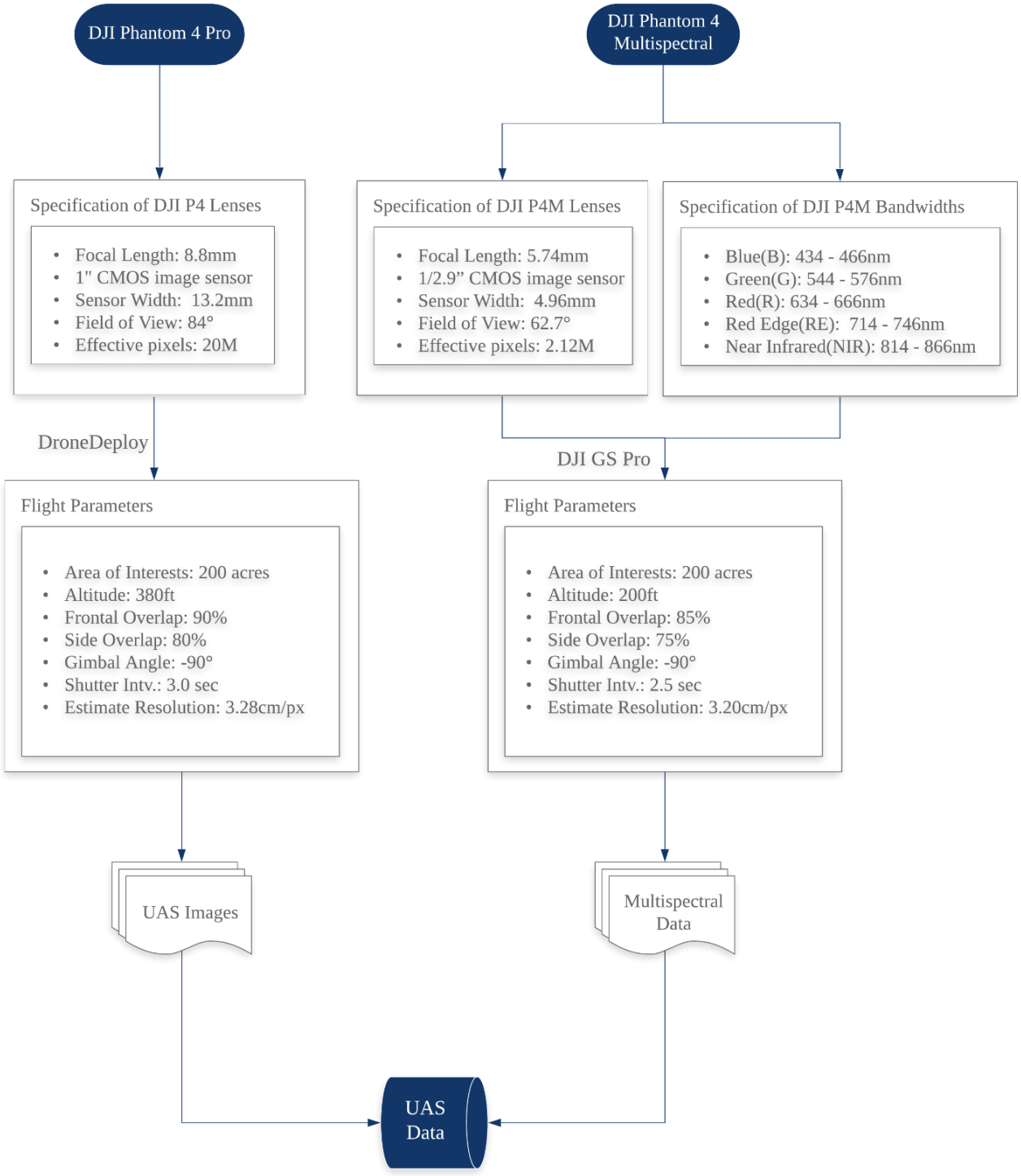

Figure 3. Data collection flowchart

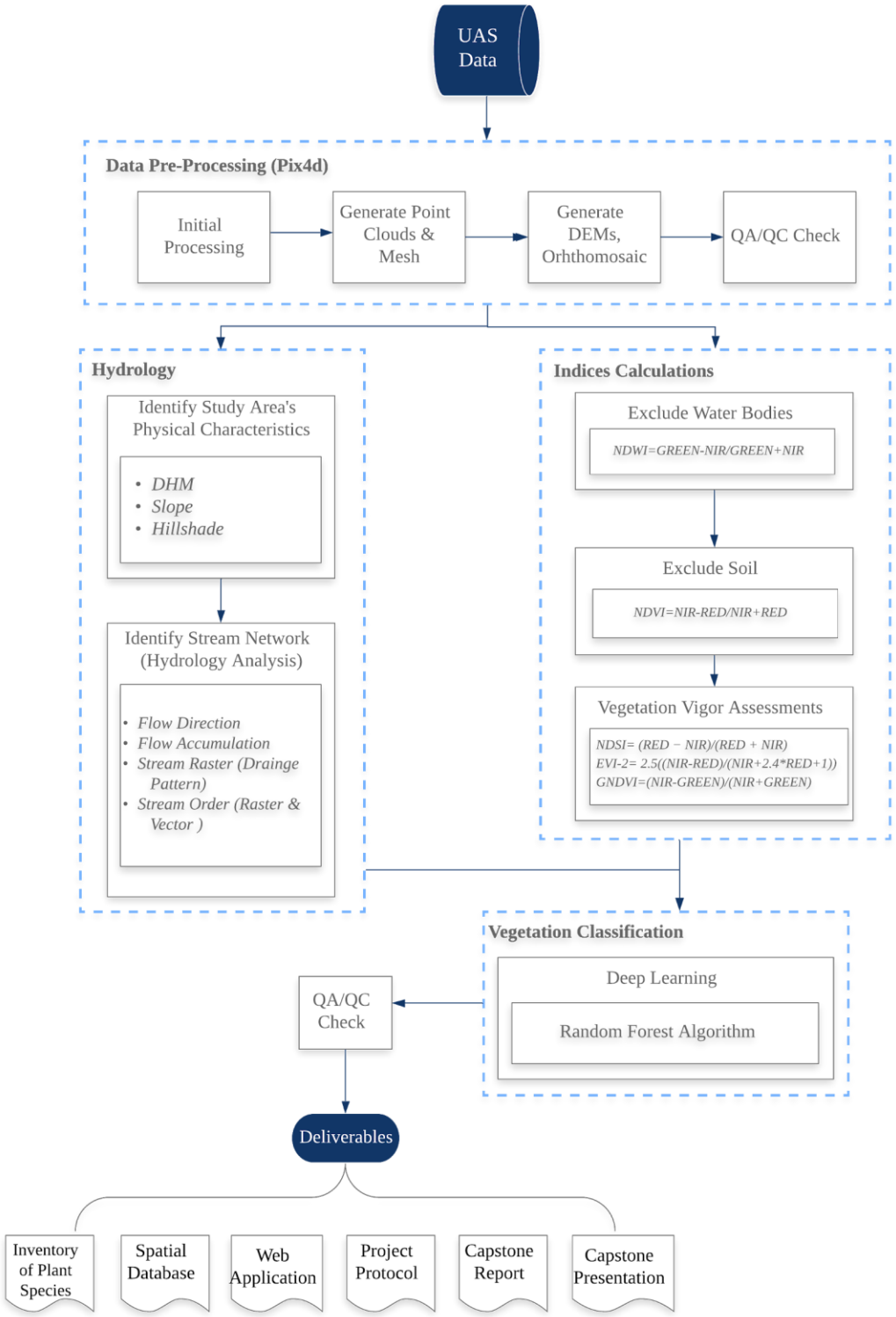

Figure 4Data processing flowchart

During this project, protocols were created so that each step could be repeated or manipulated by Green Drone AZ. One of the goals for Green Drone AZ is to develop a process that can be easily replicated to classify and inventory plant species using the drone data over multiple seasons and years.

2.2.1 Data Collection

UAS are changing the mapping/recording methods of vegetation growth and distribution. Compared to field surveys, which can take four hours to map a 0.01km^2 plot, some drones can cover 10km^2 in 45 minutes (Baena et al., 2017).

For this project, the data were collected in two phases, using a DJI Phantom 4 Pro (DJI P4P) for phase one and a DJI Phantom 4 Multispectral (DJI P4M) for phase two (Figure 5a and b). During the first phase, basic digital imagery was captured to examine the physical characteristics of the study area such as the gradient and flow accumulation patterns. Phase two was completed by capturing imagery that allowed enhanced vegetation species identification.

Figure 5. The DJI Phantom 4 with the two different cameras attached. a)The Pro camera b) the Multispectral camera.

Before the UAS flights were conducted, Ground Control Points (GCPs) were located in the field. GCPs are important for accurately georeferencing the images captured from a UAS. A total 28 GCPs were sited in this study area (Figure 6). The GCP locations were intentionally chosen based on Pix4D manual recommendations and placed in open areas to ensure visibility. The DJI Phantom 4 Pro captured RGB images only captured the red, green and blue bands of light. The flight altitude was 380ft (approximately 115.82 meters), gimbal angle perpendicular to ground surface. The imagery frontal overlaps were 90%, side overlaps were 80%, and images were captured every 3.0 seconds with an estimated image resolution of 3.28cm/pixel (Pix4D SA, 2017). Drone Deploy was used to visualize and program these flight parameters and flight plan (Figure 6). These data were collected on the same day.

Figure 6. Ground Control Points (GCPs), Drone Deploy and Pix4D flight plans

The DJI Phantom 4 Multispectral (P4M) was used to collect blue, green, red, red-edge, and NIR bandwidths. The parameters for this flight included a flight altitude of 200 ft (approximately 60.96 meters), gimbal angle perpendicular to ground surface, frontal overlap of 85%, side overlap of 75%, with images captured approximately every 2.5 seconds with an estimated image resolution of 3.20 cm/pixel (Pix4D SA, 2017). Due to limited battery life a total of 42 flights were made with the P4M. The site was also split into five 50 acres sections (as separate flight missions) with each mission plan overlapping the other. Included in the protocol is a flight planning manual that explains how these parameters were selected in more detail.

Once flights were completed, the images were uploaded to Pix4Dmapper 4.5.6. The output coordinate system for the site was set to NAD 83 / UTM zone 12. Also, an initial process was started to create key points, which are points between images that match to one another (Figure 7). After the initial process, an excel file (.csv) of the ground control points was imported and marked in the photos to improve locational accuracy. The next step was to orthorectify the data set to verify no major errors from georeferencing. After the quality check, Pix4D was used to create the densified point cloud. Point clouds take the X, Y, and Z coordinates of points and reconstruct it into a 3D model with points containing the bandwidth information. Since the DJI P4P data was processed first, this point cloud was used to generate the Digital Terrain Model (DTM), the Digital Surface Model (DSM), and the orthomosaic. Once the DJI P4M data was processed, the point cloud was further processed to create the raster datasets for index calculations. Further details on the Pix4D image processing manual are included in the protocol.

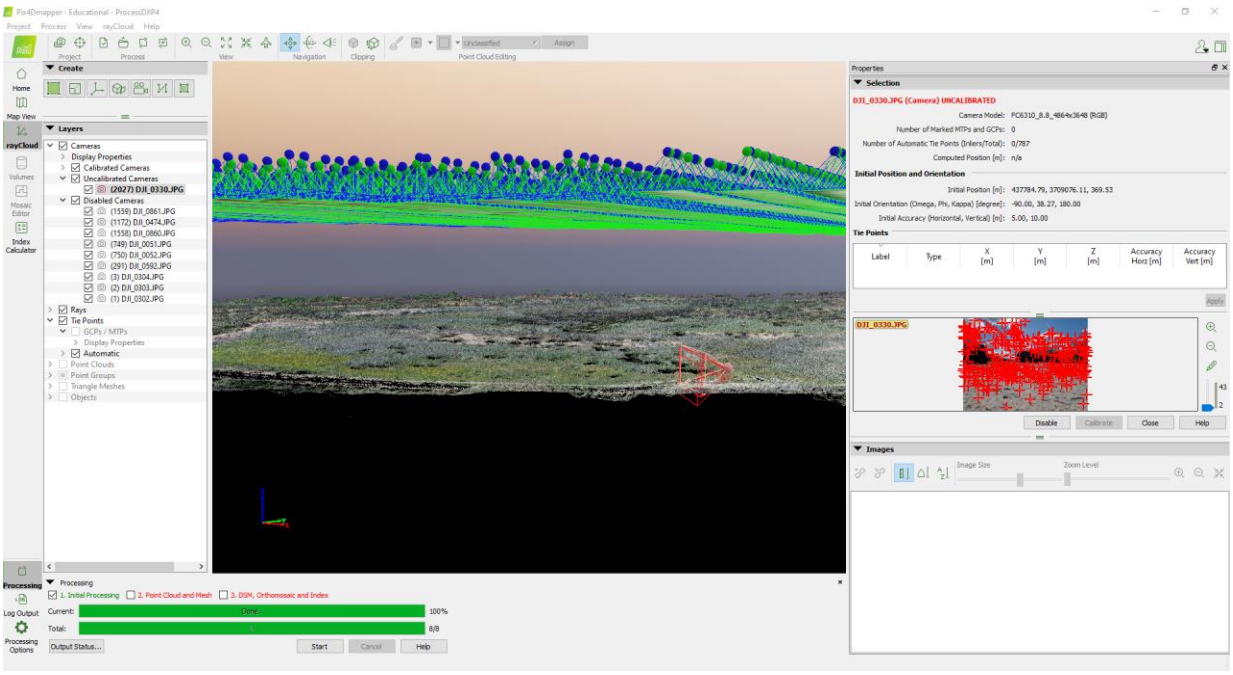

Figure 7. Showing the Pix4D application with the camera locations above their corresponding images.

Due to limited literature on using Pix4D to process massive amounts of multispectral data to classify vegetation data, there were a lot of challenges and trials with the Pix4D application. The DJI P4 flight data was much easier to process with 2,268 images and a day to complete. The DJI P4M flight had a total of 78,552 images. Pix4D recommends processing only 2000 images at a time for optimal processing time. Since the flights were split into 5 sections, due to limited batteries, processing was completed per section. Processing time was 2-3 days for each section of data It is also important to note that while processing the P4M data, the RGB images were not processed to reduce the time and memory space needed to process multispectral bands.

Once the orthomosaic was generated, it was uploaded into a web map, linked with Collector for ArcGIS and a Bad Elf GPS Receiver for field observations. Field data were collected on the current plant species and pictures were taken of multiple plant species including their coordinates. These data improved classification of the orthomosaic for deep learning algorithm training.

2.2.2 UAS-derived Terrain Data for Hydrology

There has been a growing interest among researchers and policymakers to monitor regional hydrology and anticipate potential flood hazards (Heintz et al., 2012). In order to study the physical characteristics of the study area, the terrain and surface water features were analyzed using the Hydrologic Engineering Centre – River Analysis System (HEC-RAS) and ArcGIS Pro (Esri, 2016). Flood models and surface water features were derived by characterizing UAS-derived Digital Surface Model (DSM), Digital Terrain Model (DTM), and Digital Height Model (DHM) or Canopy Height Model (CHM), which are all Digital Elevation Models (DEMs) for hydrological studies (Figure 5).

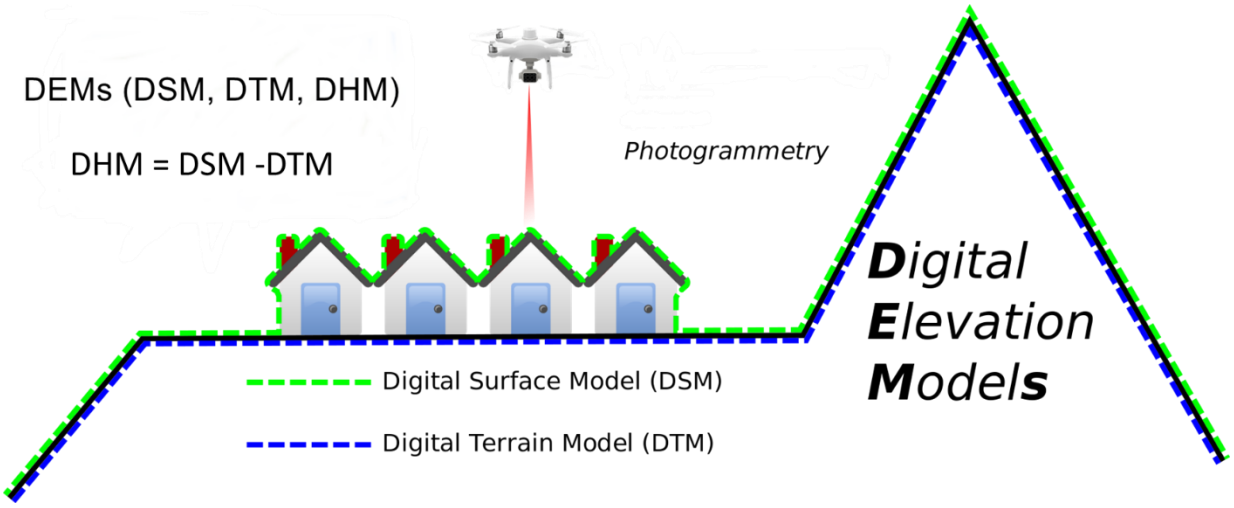

Figure 8. A comparison of the digital surface model and the digital terrain model. (modified after Arbeck - Own work, CC BY 4.0)

DTMs and DSMs are widely used to visualize and analyze the surface topography (Orengo and Petrie, 2018). The DSM enables analysis of surface obstacles or manmade objects such as bridges, roads, walls, vegetation, and other structures that may alter the flow direction of a stream, river channel or other water feature, and this process only requires the interpolation of surface elevation values (Dong and Chen, 2018).

The DTM provides crucial information about the terrain geometries such as river channel slope and flow capacity and requires one to first identify elevation values associated with the ground/terrain, and then interpolate those values into a raster surface (Dong and Chen, 2018). The DHM finds the differences between the DSM (object elevation) and the DTM (ground elevation) to represent the heights of everything (buildings, canopy, and objects such as tires, sandbags).

HEC-RAS software was used for visualizing flood simulations (analysis) using DTMs (Mokhtar et al., 2018.). It is a widely used, and well-established open channel flow analysis software tool, developed by the Hydrologic Engineering Center, an organization within the U.S. Army Corps of Engineers (USACE). HEC-RAS processes model outputs and displays them with the initially imported terrain data. River floods and precipitation are major sources of water to the study area, which controls vegetation growth and distribution.

Terrain (DTM) data derived from the UAS imagery was added to HEC-RAS. RAS Mapper was used to set the project's coordinate system (NAD 83 / UTM zone 12). The first step of setting up the hydraulic model was defining the extent of the modeling domain over which the software will compute using a 2D flow Area window. In the 2D Flow Area editor window, the default manning's n value was changed to 0.04. The Manning coefficient represents the type of surface in your modeling domain. A value of 0.04 represents light brush and trees, which is the predominant land cover in our study area (Chow, 1973). Computation point spacing was entered using a preferred mesh cell size (10x10 ft). The cell size is small enough to capture the size of landforms in the modeling domain and large enough to minimize time taken for the model to run. Force mesh recomputation in the 2D flow area editor window was used to digitize the 2D flow area generated mesh. (Figure. 9a). The SA/2D Area BC Lines tool was used to identify the location for streamflow to enter and exit the mode. The polylines are drawn along the outer boundary of the border (without

touching the boundary or crossing the boundary line) approximately the length of the width of the stream. The Unsteady Flow Model Setup tool was used to define the amount of flow to enter the model (Figure. 9b) including the Friction Slope for this are (n=0 .032) HII BARNES - pubs.usgs.gov.

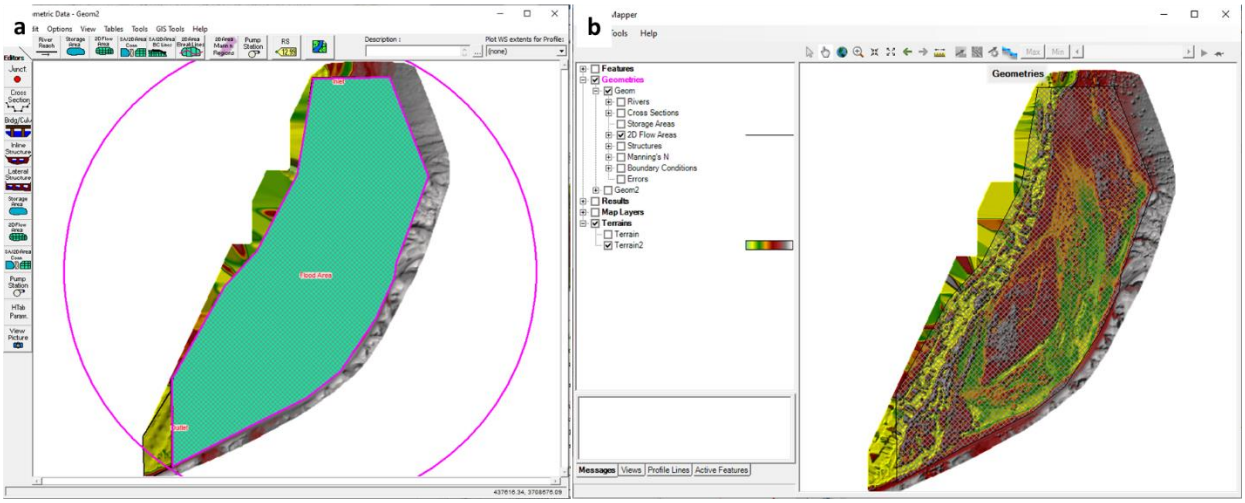

Figure 9. 2D flow area indicating inlet boundaries, outlet boundaries and mesh for flood

The Flow Hydrograph window was used to input flow data for the study area from the USGS Water Resources website. Flow data for the period of July 2019 – July 2020 was downloaded and plotted (Figure 10a). A high discharge anomaly was identified on September 23rd 2019 (Figure 10b) due to high rains and the release of the dam upstream during that period. These data were entered into the date, simulation time, and flow columns of the hydrograph data window. These data were computed at 15-minute intervals, under Unsteady Flow Simulation. The Starting Date and Ending Date for Post Processor options were entered. The starting Time was 00:00, and Ending Time was 16:00. Finally flow options were set to full momentum for better results and the Unsteady Flow Model computed.

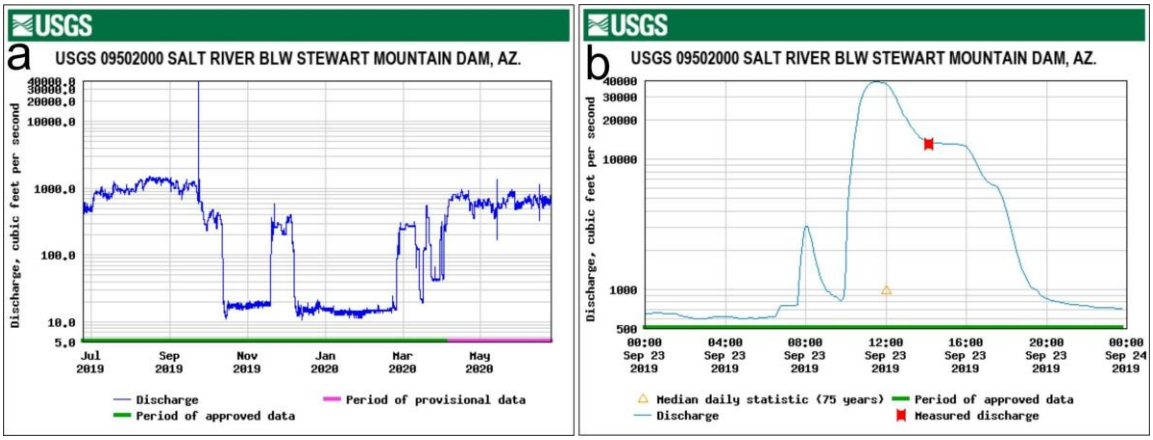

Figure 10. a) Flow data for the period of July 2019 – July 2020, b) High discharge anomaly on September 23rd 2019

The simulation results were displayed within RAS Mapper. After model execution, HEC-RAS processes the model outputs and displays them with the initially imported terrain data. There, various modeled flow properties were visualized including flow depth, velocity, and elevation. Animation of the modeled flow was viewed with particle tracing, static velocity arrows, exported and saved for interpretation.

Using ArcGIS Pro, the DTM was used to identify surface water features and to analyze slopes and hill shades to further map the study area terrain to identify flood and drainage patterns (Esri, n.d.). These data were analyzed using hydrography tools and techniques to reveal drainage patterns, channels, and water traps after precipitation or floods. These analyses were achieved using Geographic Information System (GIS) Hydrology tools in ArcGIS including:

- Flow Direction tool to determine the direction in which water would flow out of each cell.
- Sink tool to identify depressions in the DTM that are where water flows into them but cannot flow out.
- To ensure proper drainage mapping, these depressions were filled using the Fill tool.

- To create a stream network, the Flow Accumulation tool was used to calculate the number of upslope cells flowing to a location using the flow direction raster as input.
- A threshold was specified on the raster derived from the Flow Accumulation tool; the initial stage was defining the stream network system. This was accomplished using the Con tool (Map Algebra). The syntax used in Con was Stream_Raster = Con ("Flow_Accumulation">100, 1). This means all cells with more than 100 cells flowing into them will be part of the stream network for that area.
- The Stream Order tool was used to represent the order of each segment of the stream network using the Strahler techniques. This dataset represents areas where water accumulates after rains or flooding events within the floodplain. The Stream_Order raster is an important dataset for random forest classification.

2.2.3 UAS-derived Terrain Data for Indices Calculations

Based on literature review from multiple scientific papers, there are many different indices that could be helpful in classifying plant species. Each plant species has unique biophysical and biochemical makeup, that reflect wavelengths of light differently and at varying levels. The most commonly used vegetation index is the Normalized Difference Vegetation Index (NDVI). NDVI is used to highlight "greenness", but it is not suited for species discrimination. Using just NDVI or a single index for species identification has not had great success. However, combining multiple indices together, helps create a decision tree that will lead to accurate classification. The calculate raster tool in ArcGIS was used to create six indices (Table 1). A detailed manual for creating these index files is provided in the protocol.

The Normalized Difference Water Index (NDWI) is sensitive to change in water content among plants. The NDWI looks at the green and NIR wavelengths. Calculations with a value higher than

zero classify bodies of water (Nguyen et al, 2018). NDVI is sensitive to the change in vegetation health. The NDVI looks at the NIR and Red wavelengths, and calculations less than 0.3 were classified as soil (Nguyen et al, 2018).

Table 1. Indices used for plant species classification

Index	Equation
Normalized Difference Water Index (NDWI)	$\frac{GREEN - NIR}{GREEN + NIR}$
Normalized Difference Vegetation Index (NDVI)	$\frac{NIR - RED}{NIR + RED}$
Soil Adjusted Vegetation Index (SAVI)	$\left(\frac{NIR - RED}{NIR + RED}\right)(1 + L)$
Normalized Difference Salinity Index (NDSI)	$\frac{RED - NIR}{RED + NIR}$
Two Band Enhanced Vegetation Index (EVI-2)	$2.5\frac{NIR - RED}{(NIR + 2.4)(RED + 1)}$
Green Normalized Difference Vegetation Index (GNDVI)	$\frac{NIR - GREEN}{NIR + GREEN}$

Salt cedar holds salt in its tissues, so fallen leaves increase soil surface salinity making it difficult for native plants to establish (Northern Arizona Invasive Plants, n.d.). Therefore, having indices to examine the soil salinity will help identify salt cedar clusters. The Normalized Difference Salinity Index (NDSI) uses the red and NIR bands to highlight soil salinity levels (Khan, et al., 2005, Kim-Anh & Thanh-Hung, 2020). The Soil Adjusted Vegetation Index (SAVI) also uses the NIR and red bands and is similar to the NDVI but is better used in areas of low vegetation coverage. With these three (3) combined, a high NDSI value with low NDVI or low SAVI value, will imply a higher soil salinity with less greenness and should classify as salt cedar (Kim-Anh & Thanh-

Hung, 2020). Considering the fact that salt cedar is closely associated with cottonwood and giant reed which are not salty.

Lastly, an evaluation of greenness was developed to simulate the real vegetative cover within the LSR project site. The Two Band Enhanced Vegetation Index (EVI-2) uses the NIR and red bands, and higher EVI-2 values indicates higher plant vigor (Santos et al., 2014). The Green Normalized Difference Vegetation Index (GNDVI) uses the NIR and green bands and has been successful at measuring the varying plant greenness levels (Wahab et al., 2018).

2.2.4 Deep Learning for Vegetation Species Classification

Chapter 3

3.1 Results and Interpretation

3.1.Data Collection

A total of 2,268 images were collected with the DJI Phantom 4 Pro (P4P) and 78,552 by the DJI Phantom 4 Multispectral (P4M). These data were processed to generate an orthomosaic with a resolution of 3.28 cm/pixel orthomosaic (Figure. 12).

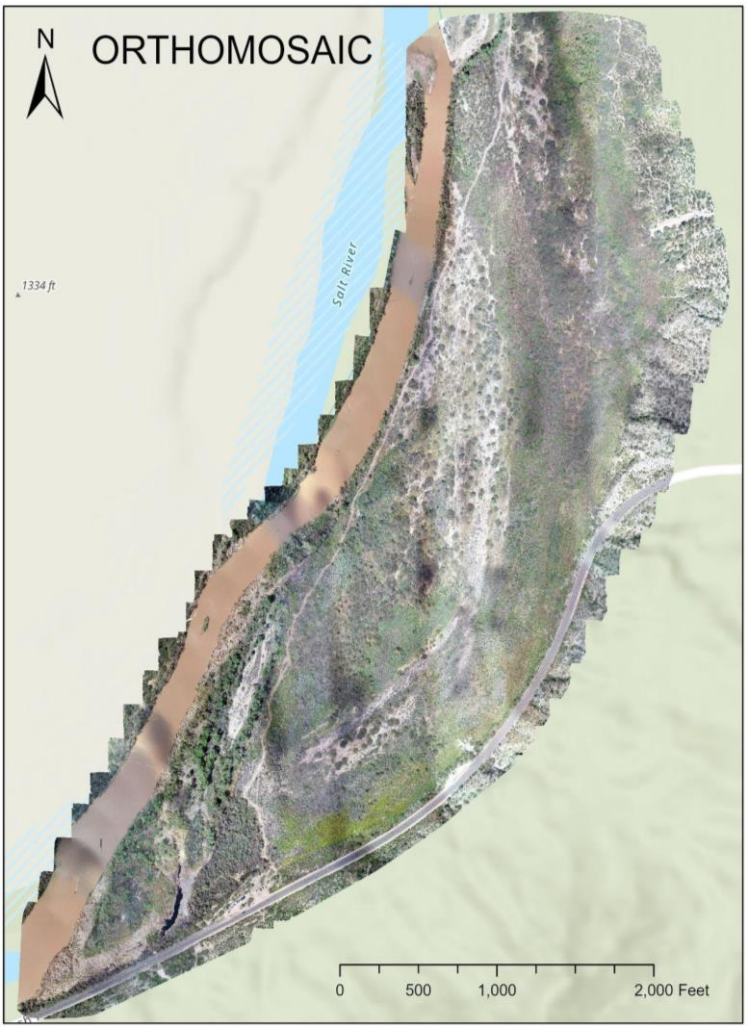

Figure 12. The orthomosaic generated from Pix4D

From field observation, (Figure 13), the project focused on eight (8) plants. These were arrow weed (*Pluchea sericea*), fremont cottonwood (*Populus fremontii*), giant reed (*Arundo donax*), globe chamomile (*Oncosiphon piluliferum*), sahara mustard (*Brassica tournefforti*), salt cedar (*Tamarix spp.*), southern cattail (*Typha domingensis*), and velvet mesquite (*Prosopis velutina*) (Figure 14).

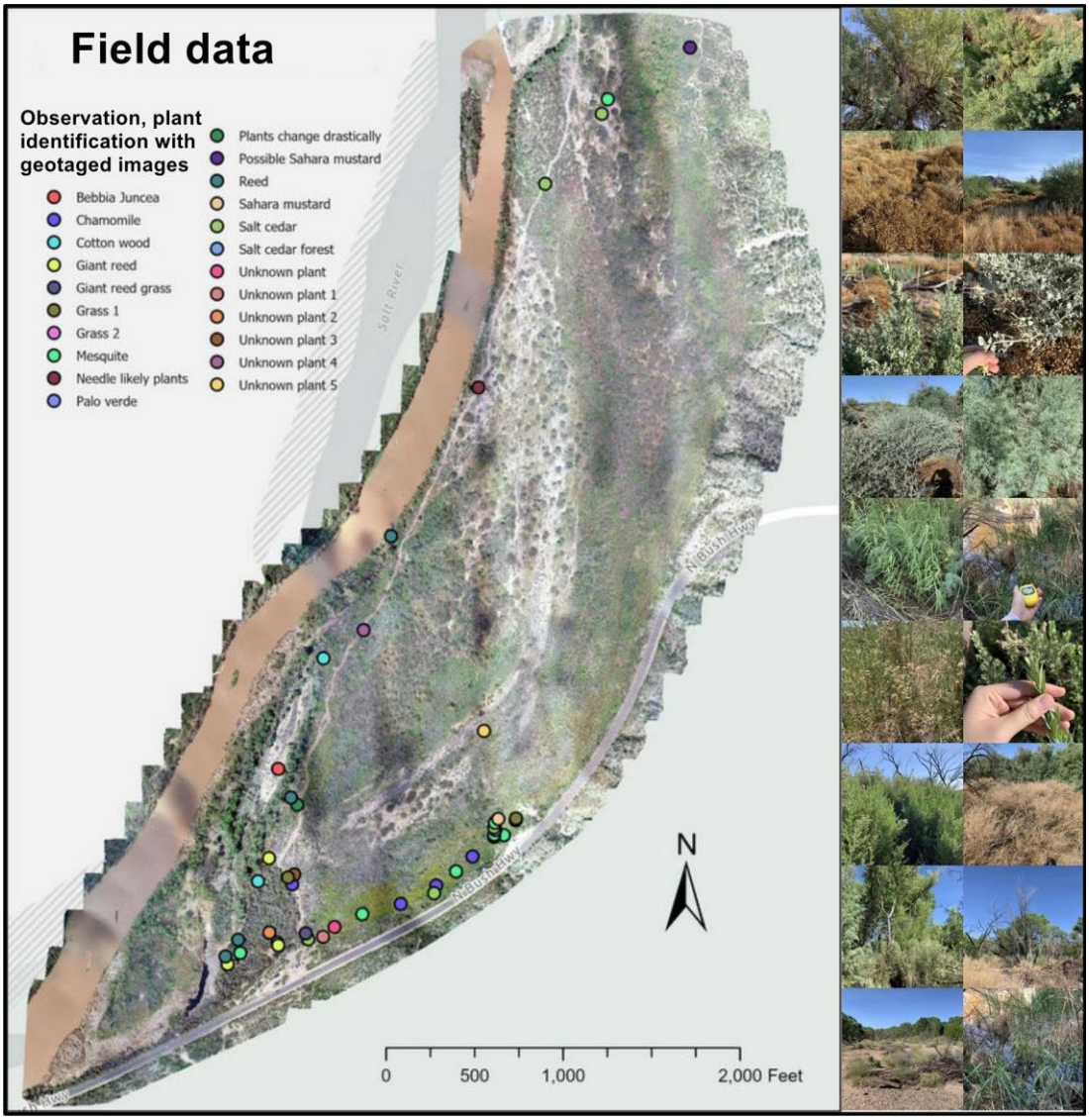

Figure 13. Display of field data on Orthomosaic to ease species signature identification

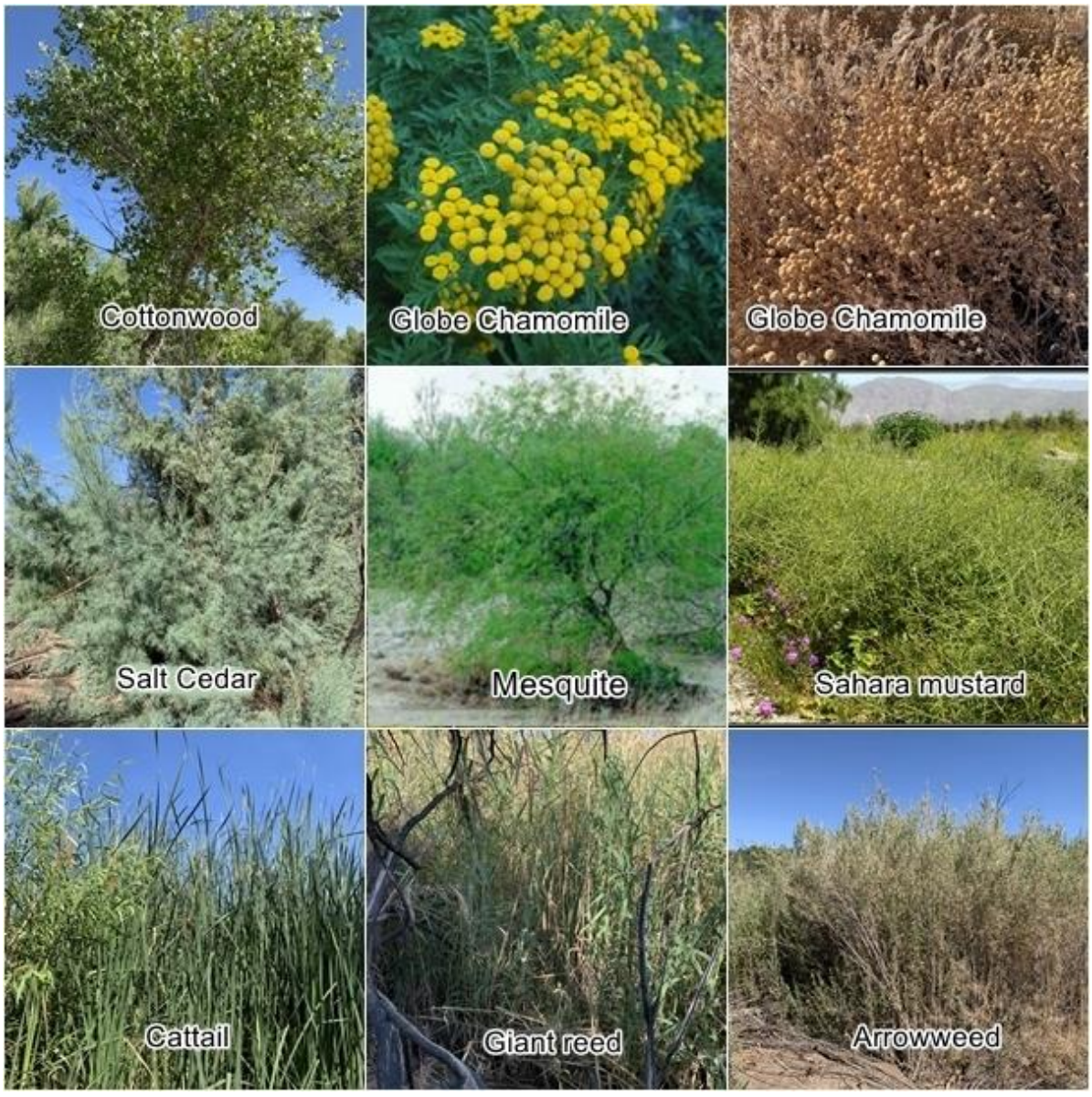

Figure 14. Eight plants (native and invasive) of interest

Using the orthomosaic and the known locations of the eight plants in the field, polygons in the shape of rectangles were used to display different plant species along with the soil, water and roads (Figure. 15). Knowing the distribution of the plant species within the study area of both invasive and native species allowed us to classify them based on their relationship to topography, hydrology, and indices. These data will be used to train and validate the deep learning algorithm.

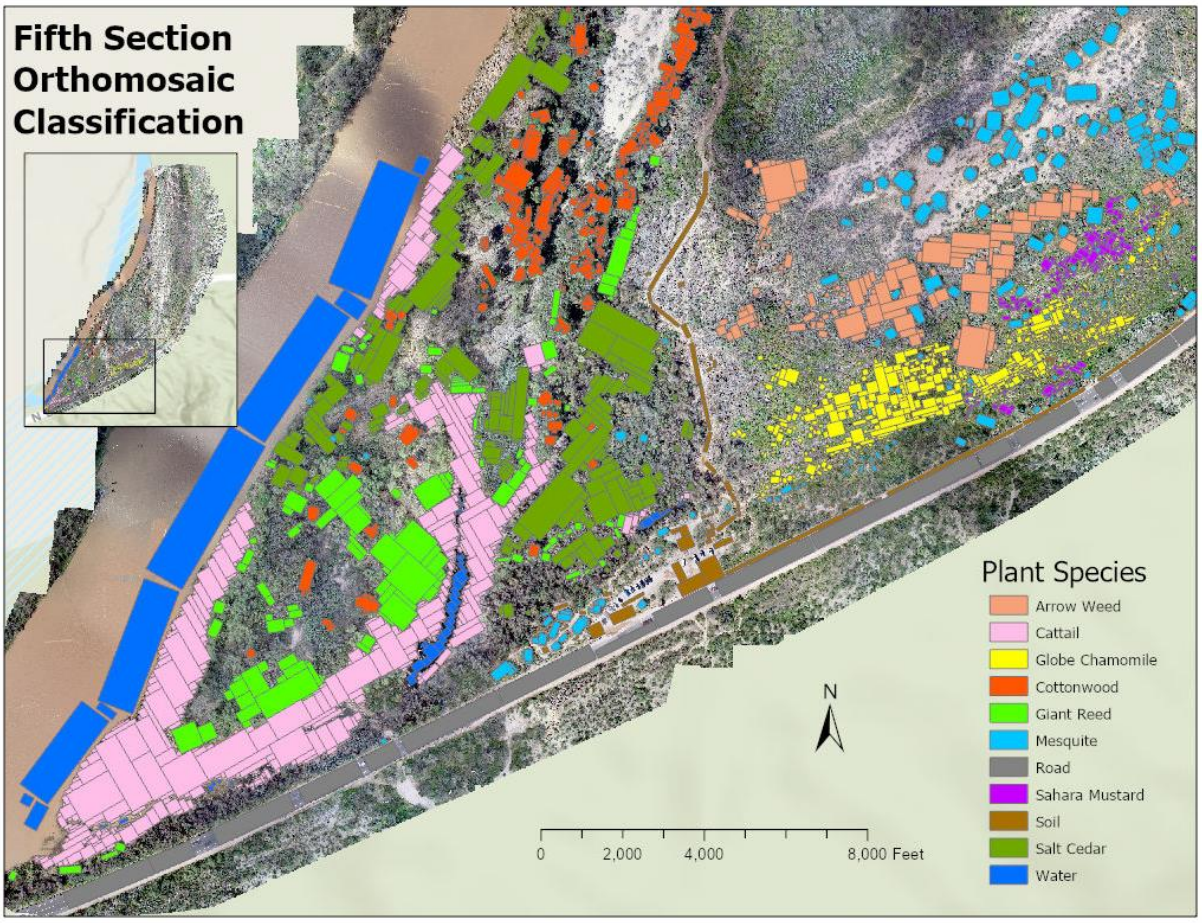

Figure 15. Final plant species classification of the orthomosaic

A Digital Surface Model (DSM) (Figure. 16) displaying topography including surface features and canopy height. Digital Terrain Model (DTM) (Figure. 17), displaying the topography without surface features included. Digital Height Model (DHM)/Canopy Height Model (CHM) (Figure 18) was generated displaying canopy height. These data ranged from about 319 feet to 389 feet of elevation and were used for hydrology and RF Analysis.

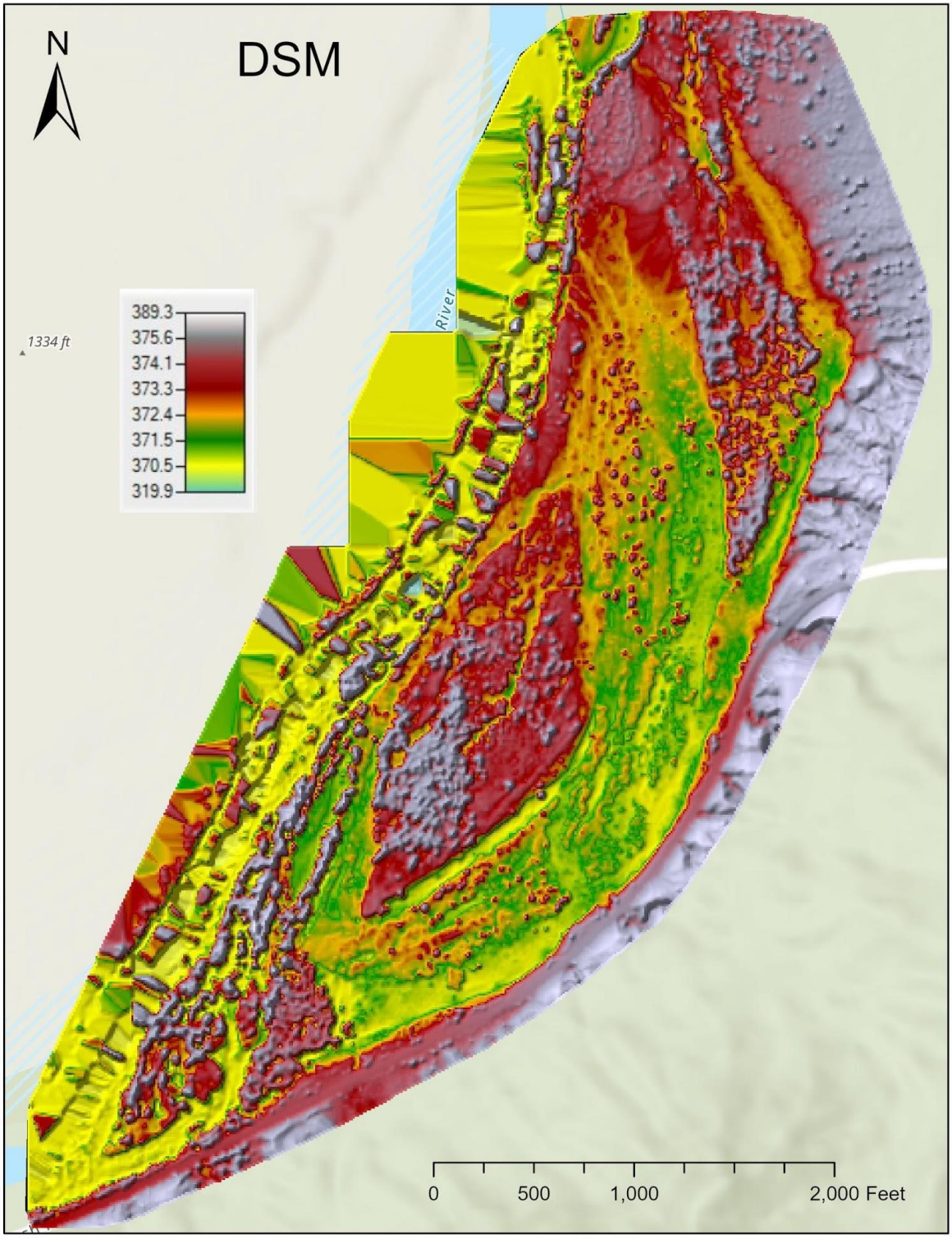

Figure 16 Digital Surface Model (DSM)

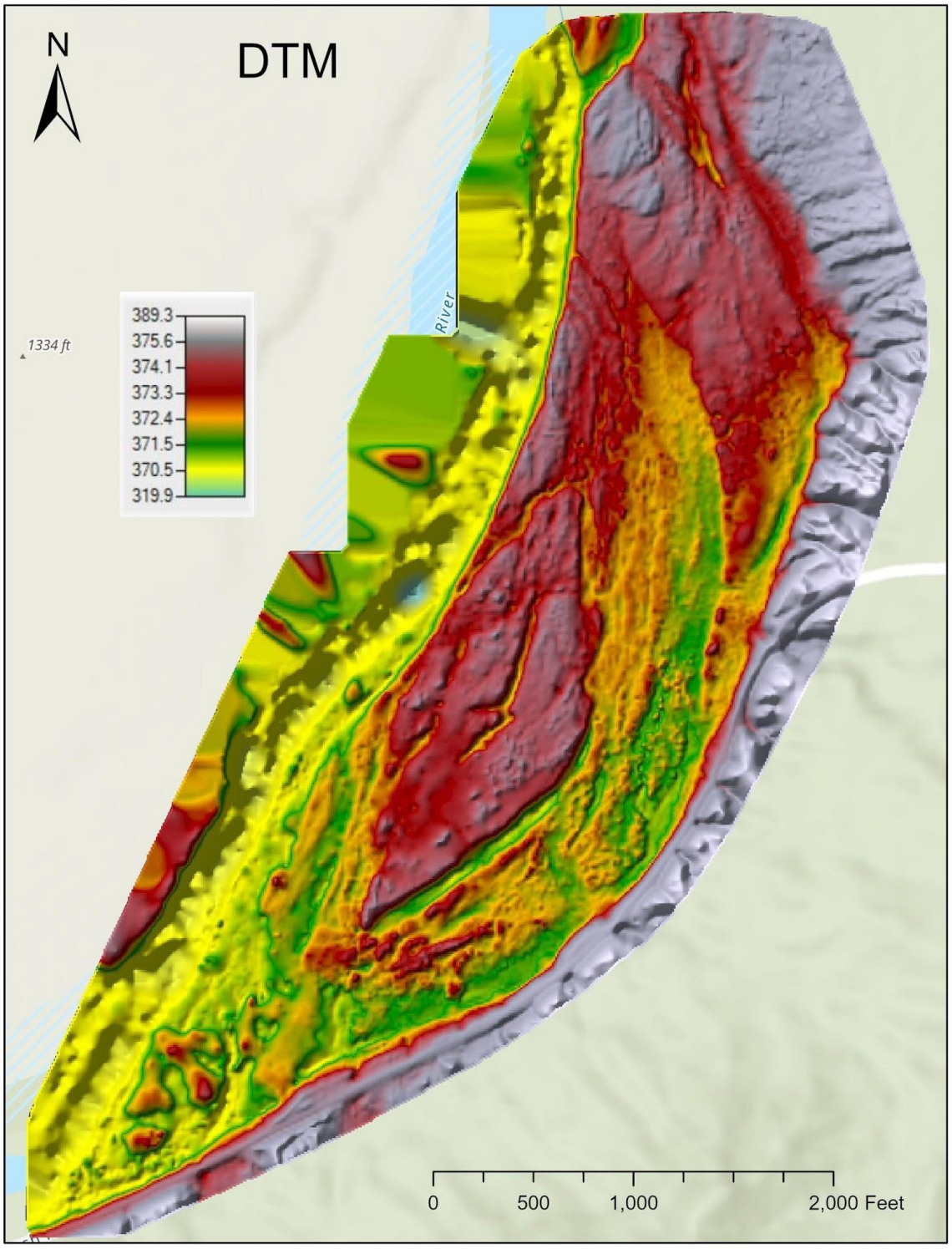

Figure 17 Digital Terrain Model (DTM)

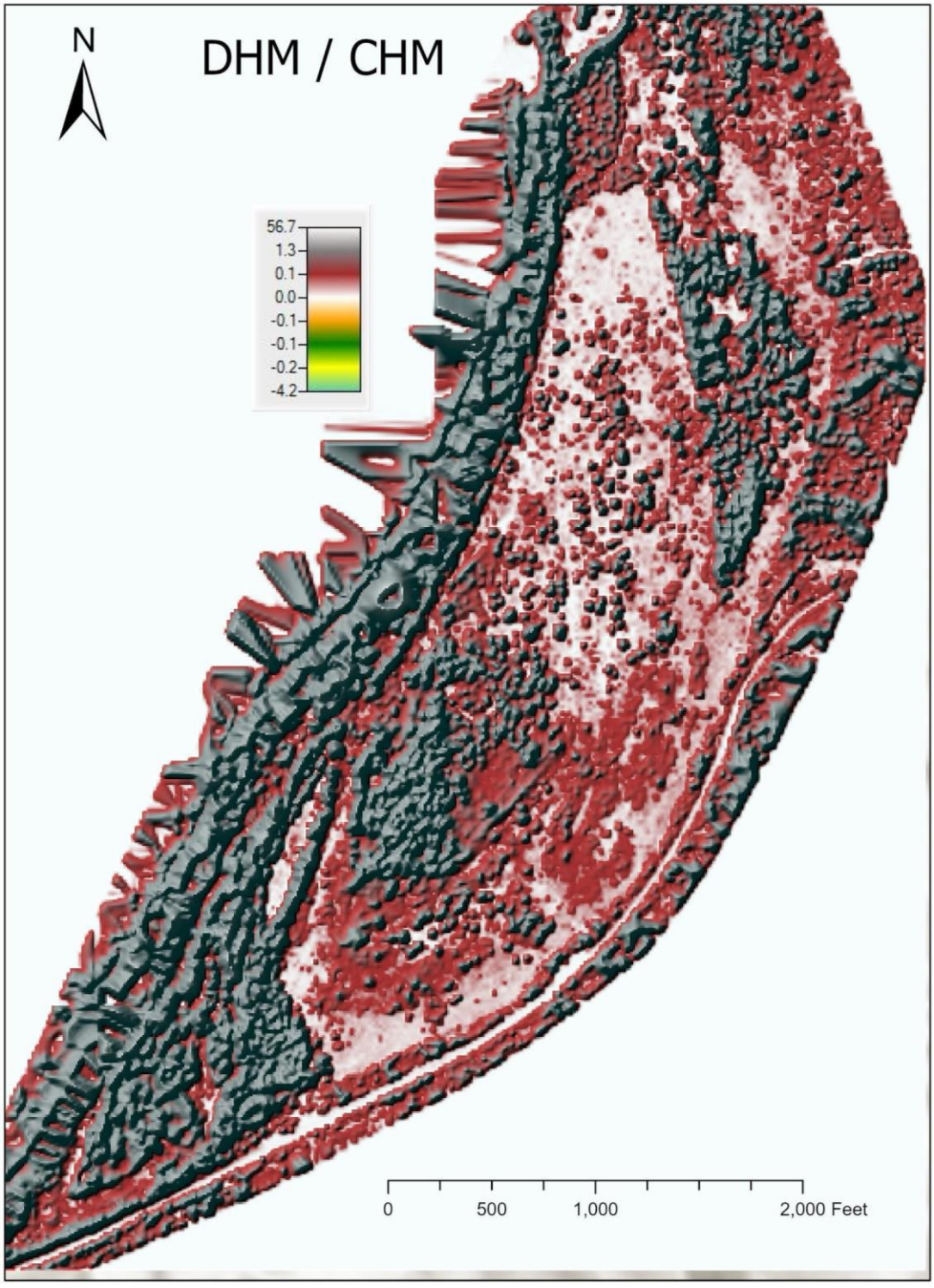

Figure 18. Digital Height Model (DHM) / Canopy Height Model (CHM)

3.1.2 Flood models and Hydrology

HEC-RAS software was used to visualize flood simulations (analysis) using the USGS flow data anomaly of Sep 23rd, 2019 for 16 hours at four different hydrograph peaks. Flood simulation for water depth within the study area (Figure 19) displays how depth changes at different flow rates within the study area ranging from 0 feet - 61 feet. These data can be used to map out hazardous areas to avoid during a flash flood.

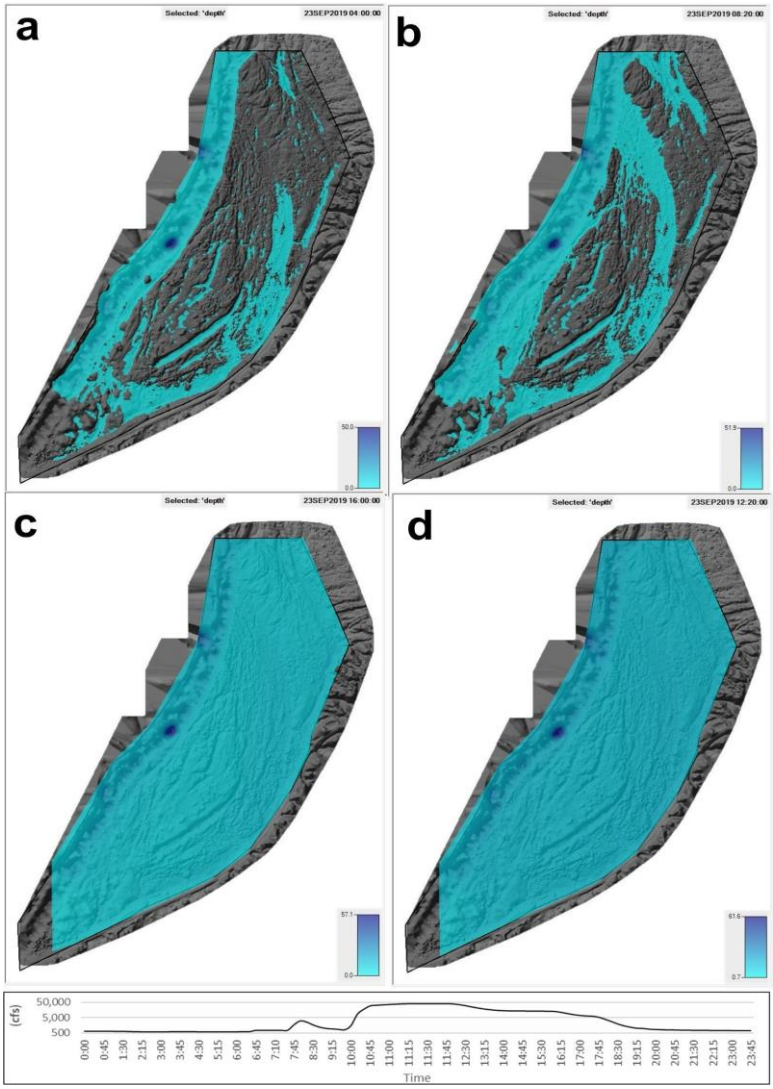

Figure 19. Flood simulation of depth ranging from 0-61.6 ft.

Flood simulation for water elevation (Figure 20) within the study ranges from 369 to 380.6 feet. These results can be used to identify areas suitable to save nursed native plants during treatment and planting periods within the floodplain where they won't be covered or destroyed by regular flash floods. These can also be used to navigate rescue attempts within the study area during a flood.

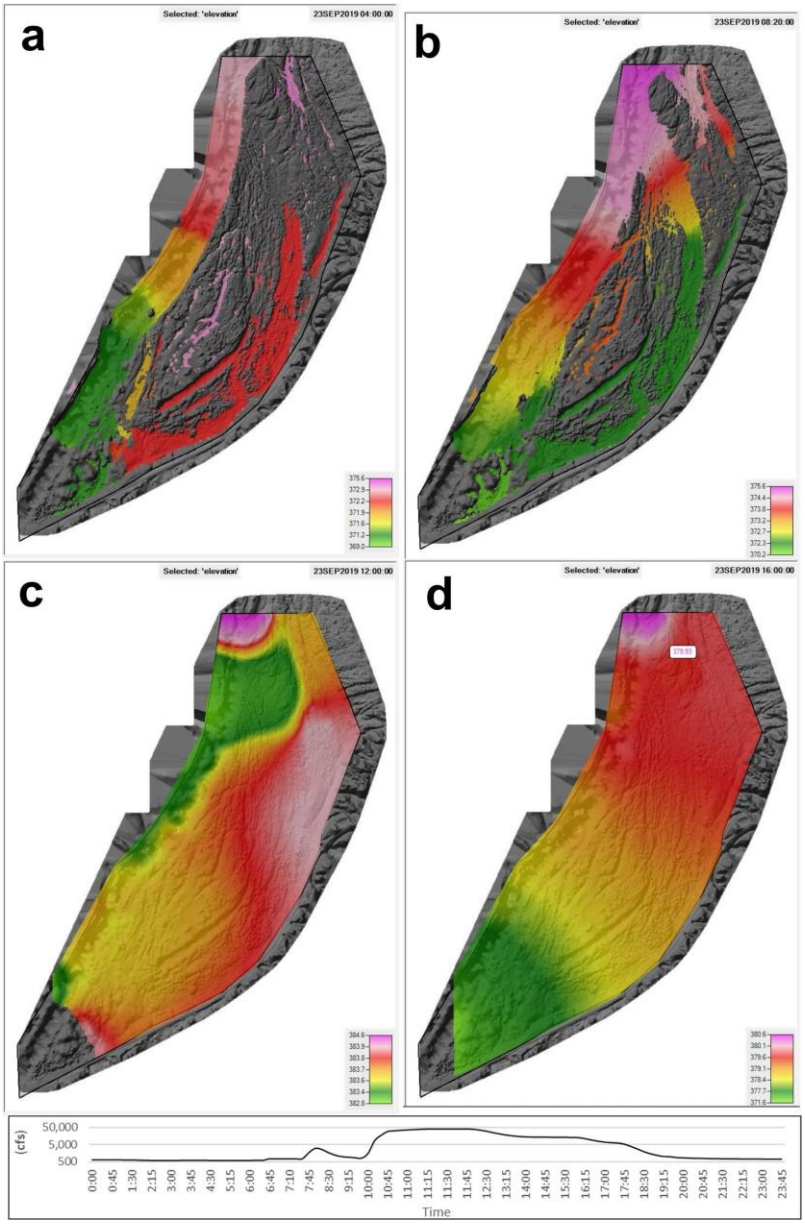

Figure 20. Water elevation flood simulation within the study ranging from 369 to 380.6 ft.

The third simulation parameter is surface water velocity within the floodplain. (Figure 21). These data identify low energy areas in blue, suitable for plant growth, and high energy areas less suitable for plant growth. The surface water velocity during floods within the study area range from 0 to 96 ft/m in some areas and helps identify areas suitable to replant native species.

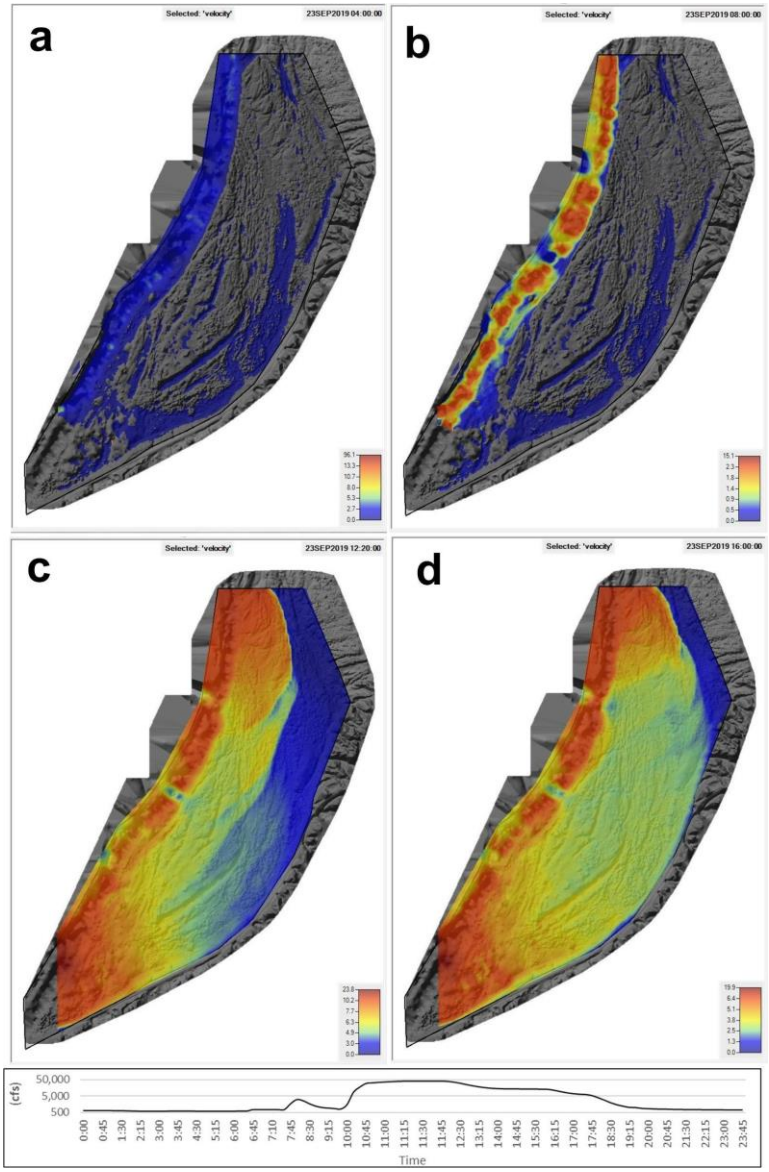

Figure 21. Surface water velocity during floods within the study area ranging from 0 to 96 ft/m.

Surface water velocity is responsible for erosion, transportation and deposition during flow. The Particle tracing tool was used to monitor particle distribution at different velocity rates within the

study area (Figure 22). This principle applies to plant seeds within the water body. Deposition only takes place in the low energy blue zones of low velocity making the southeastern area most suitable for the deposition of new seeds and growth. The river channel is a high energy environment and will only deposit particles during normal flow conditions.

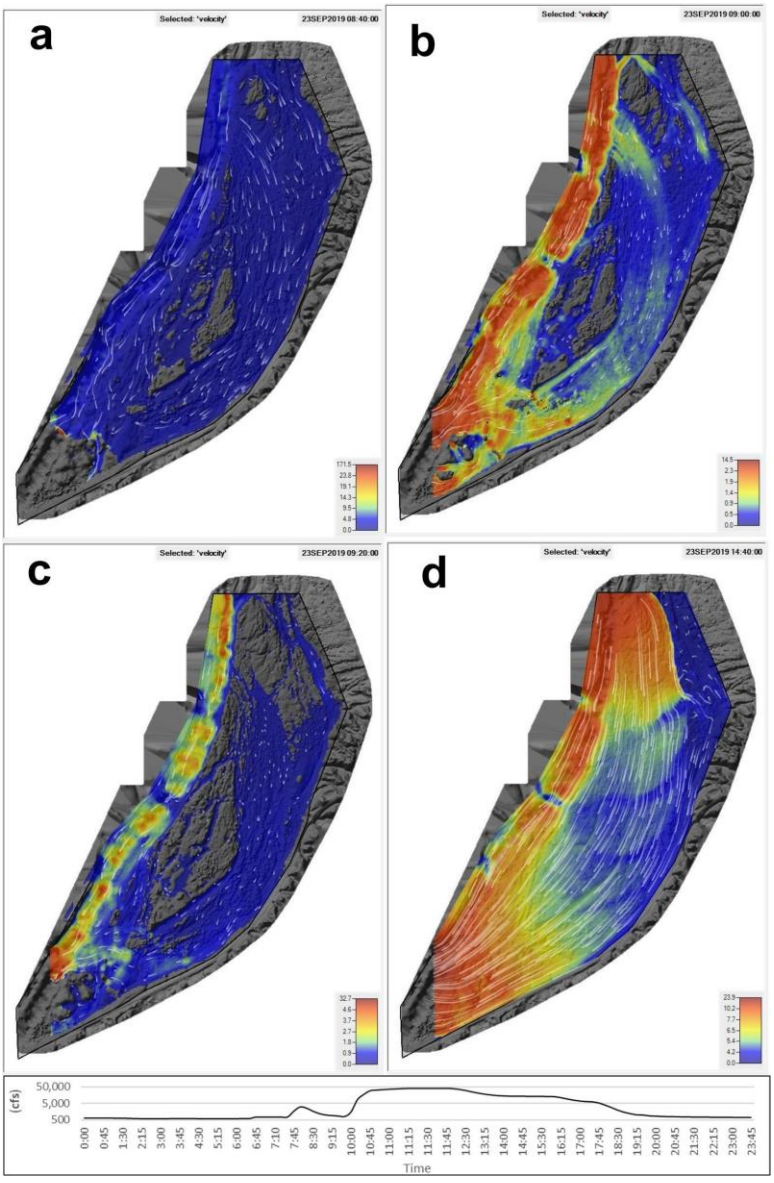

Figure 22. Surface water particle tracing during floods within the study over time within the study area.

After flooding and precipitation, water in flat areas flow away from higher ground and towards lower ground. Using the method of Garbrecht and Martz (1997) the D8 flow direction algorithm was applied to the DTM to obtain flow directions within the study area (Figure 23).

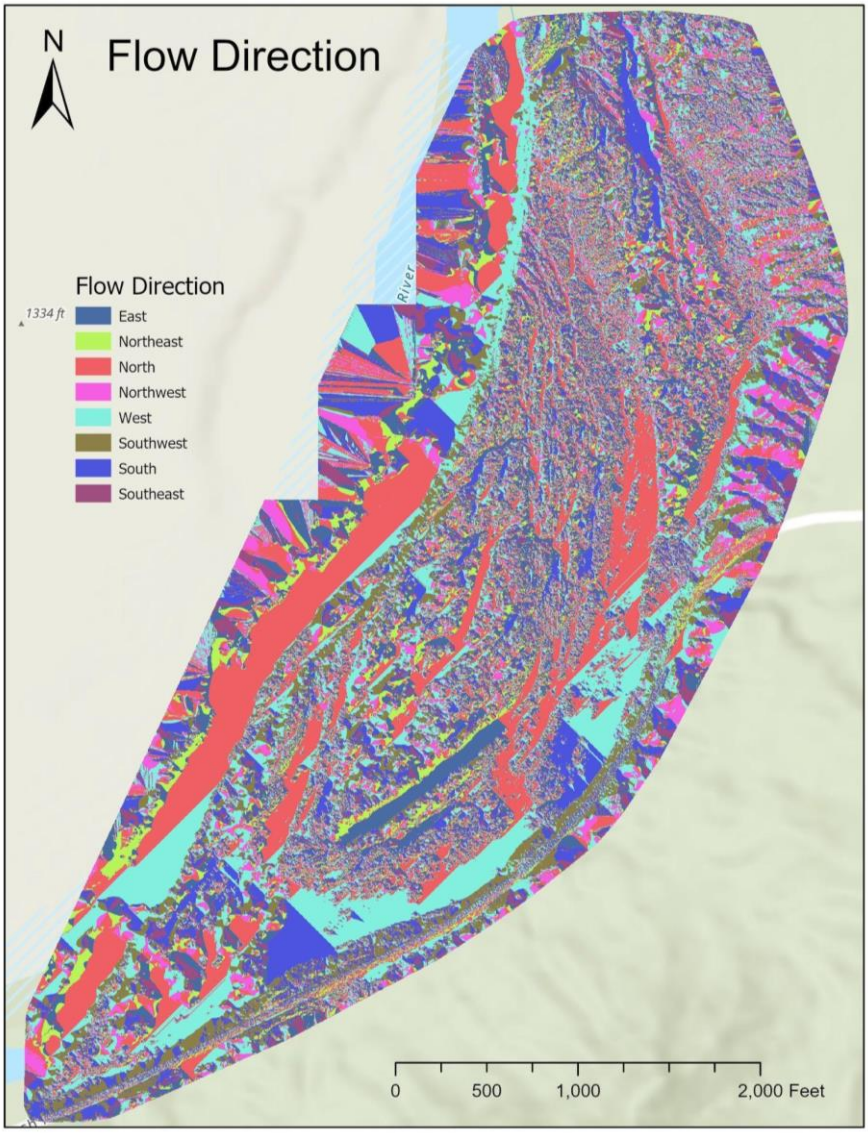

Figure 23. Flow directions of water based on gradient within the study area

These direction data were further used to obtain flow accumulation patterns within the study area (Figure 24). The Hydrology map shows steam order data which indicates potential micro aquifers where plants extract water for growth during hot and dry seasons.

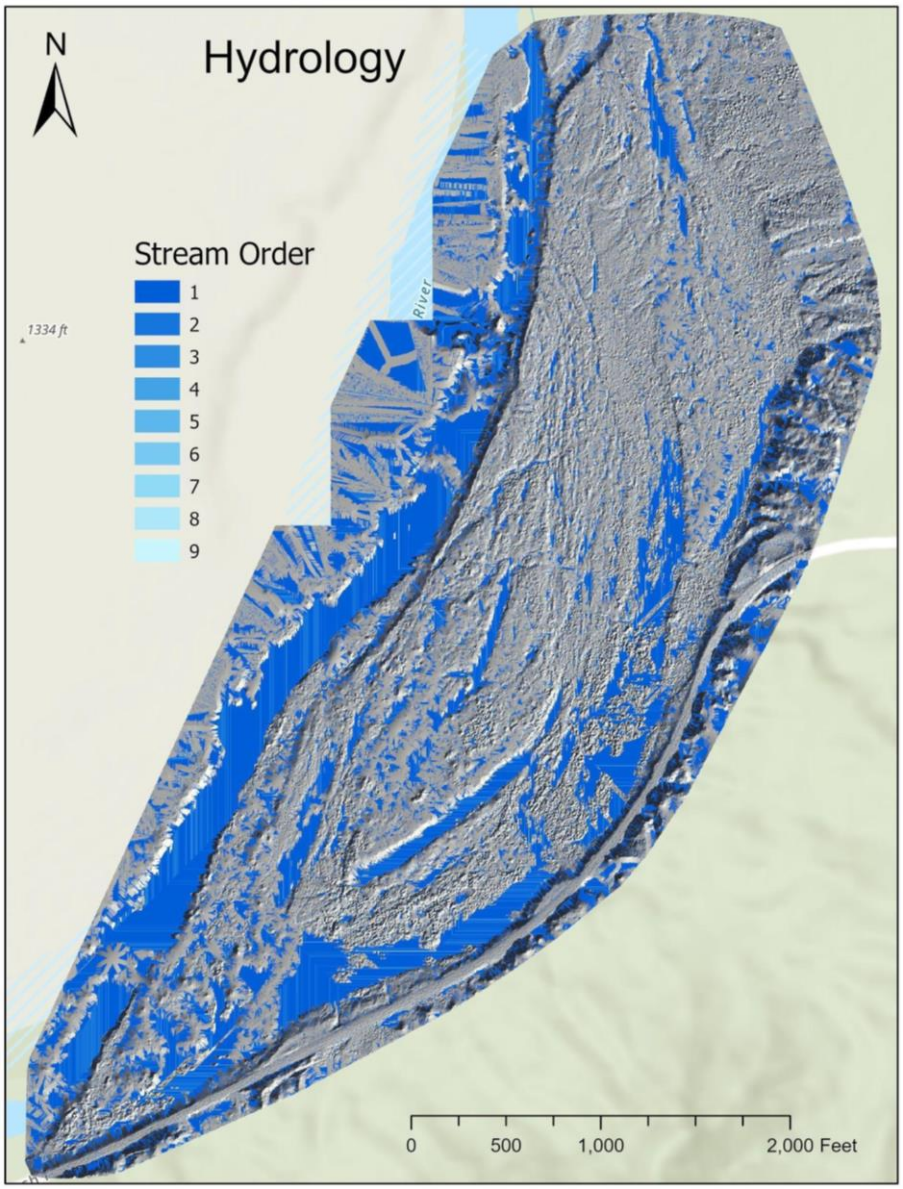

Figure 24. Hydrology map displaying steam order from flow accumulation data which indicates potential micro aquifers.

To better visualize the relationship between vegetation and hydrology, stream order and orthomosaic layers were superimposed (Figure 25). Apart from the active river channel buffer, vegetation growth seems to cluster around micro aquifers that are exposed in some areas (Figure

25b and c). Also, these areas are greener or have higher vegetation canopy heights relative to other areas.

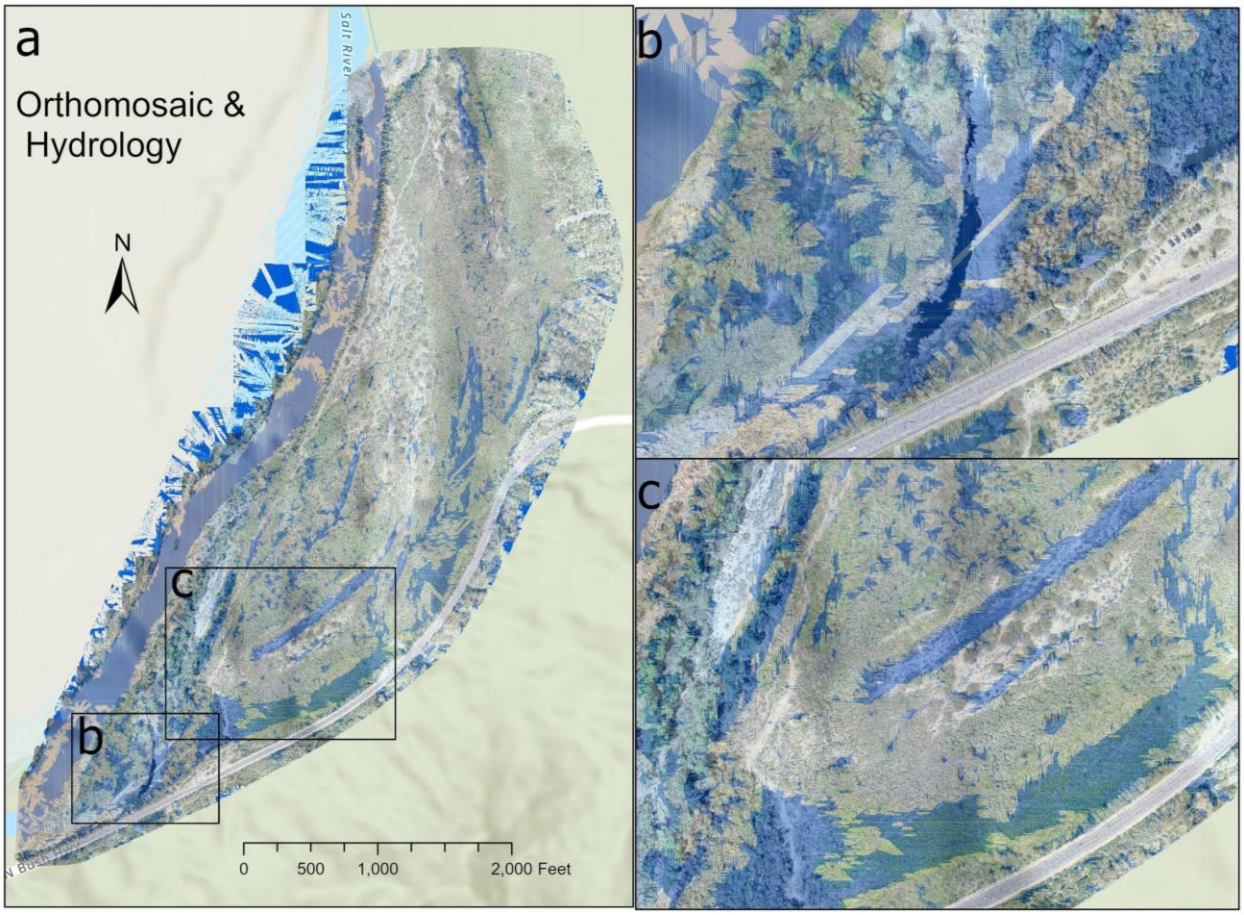

Figure 25. Relationship between vegetation and hydrology, a) Steam order and othomosaic layers superimposed, b) Microaquifer exposed to the surface, c) vegetation clusters around micro aquifers.

3.2.3 UAS-derived Terrain Data for Indices Calculations

Vegetation indices were derived for Normalized Difference Water Index (NDWI) displaying change in water content among plants with values ranging from -0.91 to 0.93 and values higher than zero were classified as water (Figure 26a). Normalized Difference Vegetation Index (NDVI) highlights "greenness"and change in vegetation health within the project site. Its values range from -0.92 to 0.93 and values less than 0.3 were classified as soil (Figure 26b).

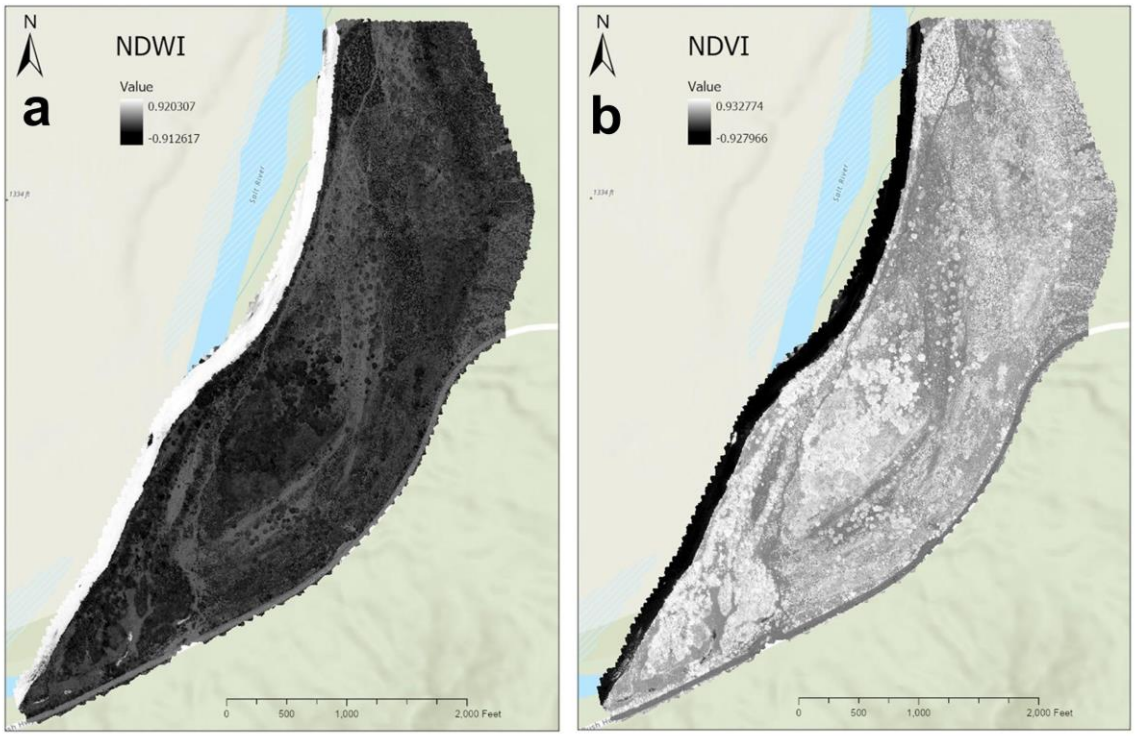

Figure 26. Normalized Difference Water Index (NDWI) and Normalized Difference Vegetation Index (NDVI)

Normalized Difference Salinity Index (NDSI) highlights soil salinity levels within the project site (Figure 27a). These salinity patterns range from -0.93 to 0.927 and also highlight salt cedars signatures within the area. Soil Adjusted Vegetation Index (SAVI) were also generated ranging from -1.39 to 1.39 (Figure 27b). Two Band Enhanced Vegetation Index (EVI-2) (Figure 28a) range from -0.003 to 0.009. Higher EVI-2 values indicate higher plant vigor in the area. The Green Normalized Difference Vegetation Index (GNDVI) (Figure 28b) displays varying plant greenness levels between -0.92 and 0.91. These indices were used to enrich the random forest algorithm.

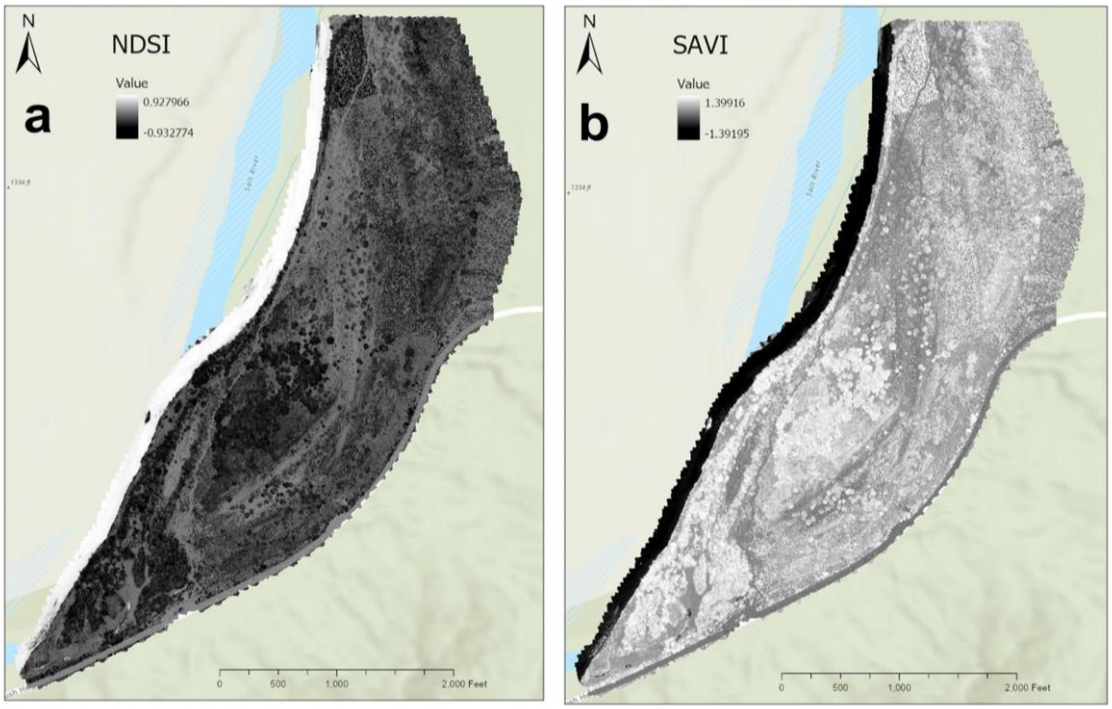

Figure 27. Normalized Difference Salinity Index (NDSI) and Soil Adjusted Vegetation Index (SAVI)

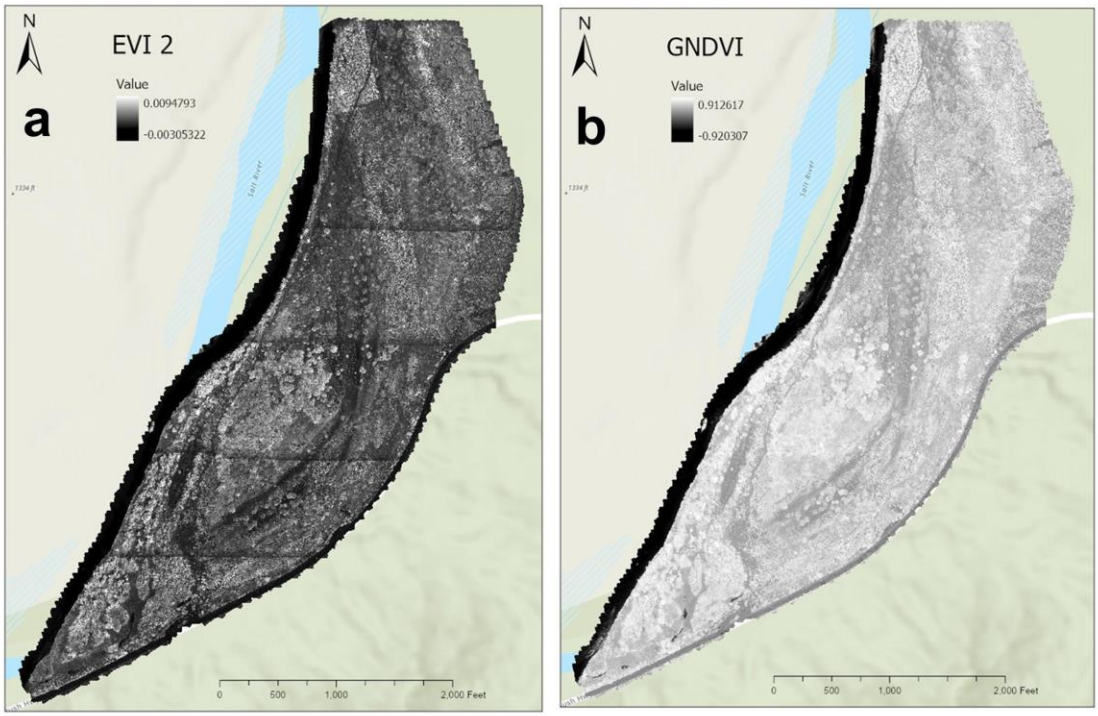

Figure 28. Two Band Enhanced Vegetation Index (EVI-2) and Normalized Difference Vegetation Index (GNDVI)

3.2.4 Deep Learning for Vegetation Species Classification

The Random Forest Model has a total of 164,237,700 pixels. Among these pixels, 24,007,620 pixels (70%) were used for training the Random Forest classifier, and 10,288,990 pixels (30%) were used for testing. This model recorded an accuracy of 89.09%, Mean Squared Error (MSE) of 2.14, and Mean Absolute Percent Error (MAPE) of 10.91%. After running the Random Forest, the feature importance in Python delivered the 10 features importance distribution (Figure 29).

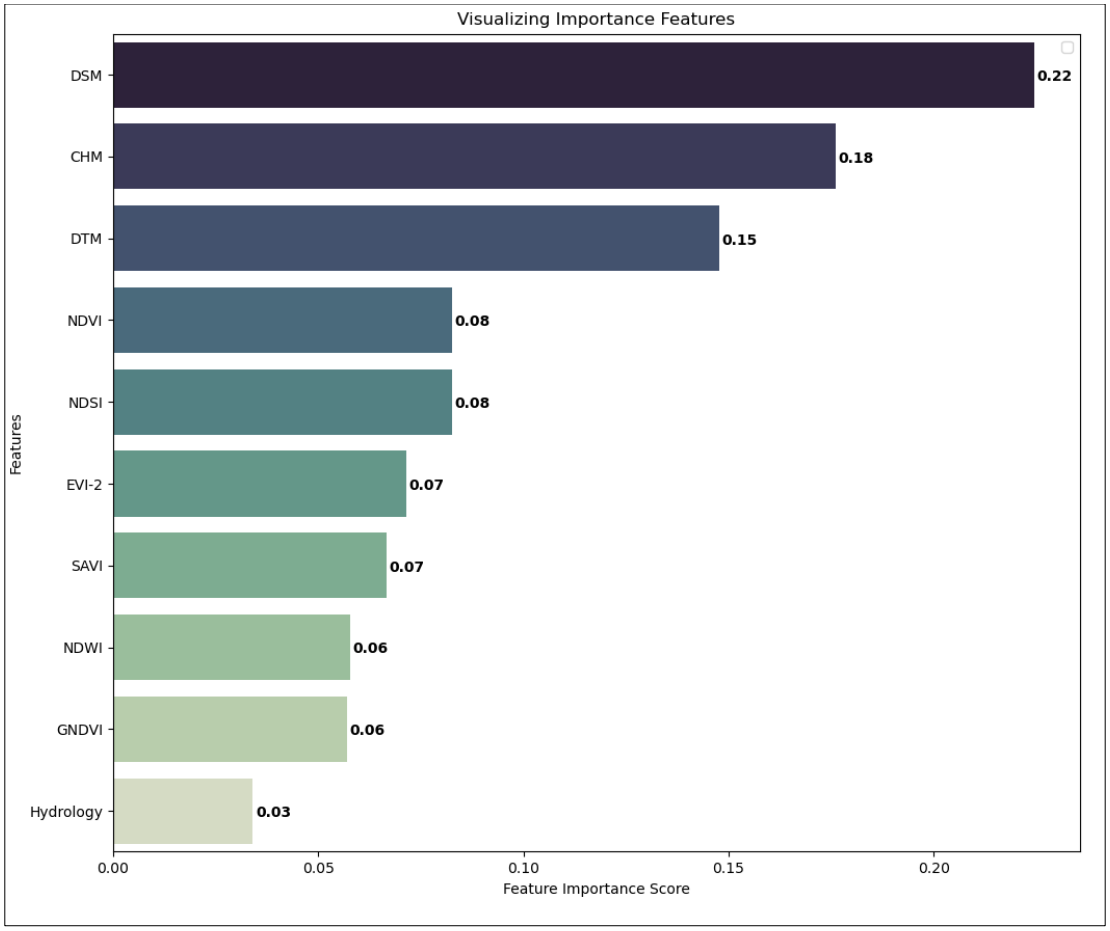

Figure 29. Lists of Feature Importance

The three (3) most significant features are DSM, CHM, DTM, which are all DEMs. These feature importances indicate that the eight (8) plants' heights are the most distinct characteristic when classifying vegetations. From the SEINet Portal Network, arrow weeds are usually 1.5-3m tall, cattails are usually 1-3m tall, cottonwoods are up to 30m tall, globe chamomiles are usually 0.5m

tall, giant reeds are usually 2-5m tall, sahara mustards are usually 0.3-1.2m tall, mesquites are usually up to 17m tall and salt cedars are usually 4-5m tall but can grow up to 8m. (Table 3) (SEINet, n.d.).

Table 3. Inches and meters chart of vegetations

Plant Name	Size in feet	Size in meters
Arrow Weed	4.92-9.84	1.5-3
Cattail	3.28-9.84	1-3
Cottonwood	up to 98.42	up to 30
Chamomile	1.64	0.5
Giant Reed	6.56-16.40	2-5
Sahara Mustard	0.98-3.94	0.3-1.2
Mesquite	up to 55.77	up to 17
Salt Cedar	usually 13.12-16.40, up to 26.25	usually 4-5, up to 8

One of the 200 decision trees in the Random Forest was visualized (Figure 30). This decision tree has a total of 384,010 decision nodes, which is too large for visualizing. Therefore, Figure 30 shows the first 15 decision nodes of the decision tree. After running the Python codes for the Random Forest algorithm, a classification raster image was generated for each plant species (Figure 31). The percentage of these species will be calculated and displayed on a dashboard.

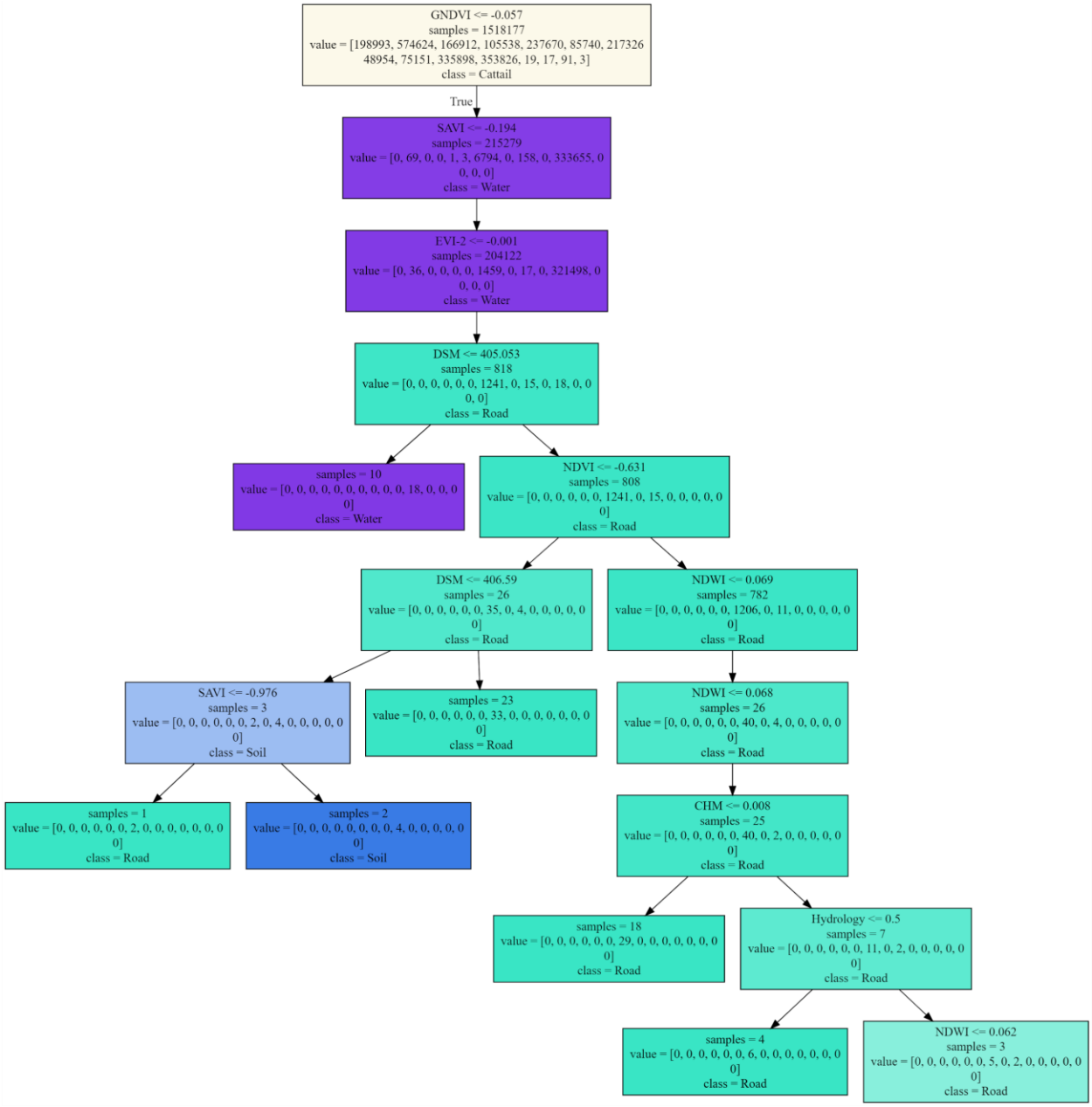

Figure 30. First 15 nodes of a decision tree from the Random Forest Algorithm

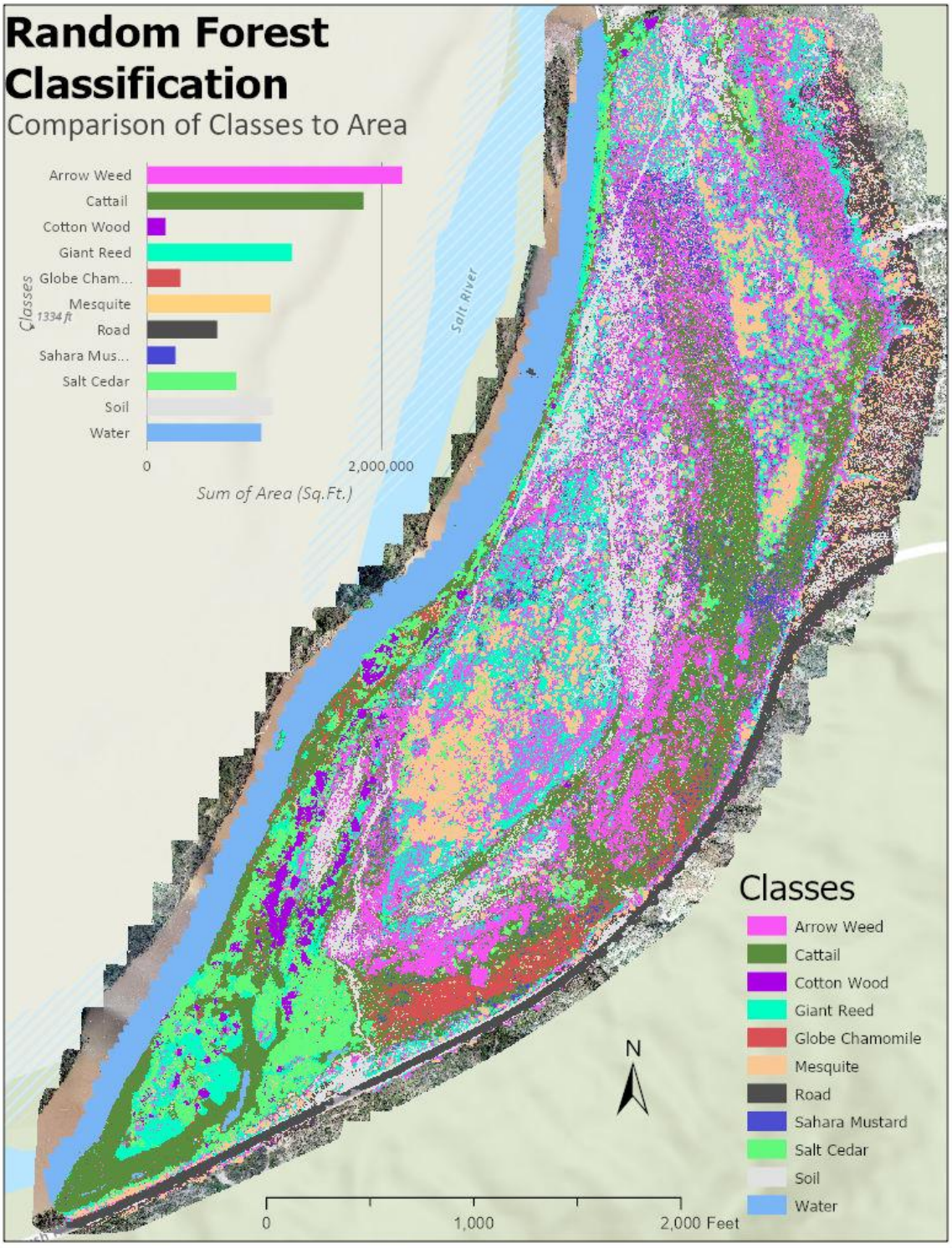

Figure 31. Classification raster image generated for each plant species

Chapter 4

4.1 Discussion and Conclusion

For the preprocessing and data collection phase, it is recommended to have larger and more defined ground control points to allow higher accuracy when marking them in Pix4D. These ground control points should be placed and/or inspected before each flight. This would also be a good time to mark plant locations for the field collection of plant species locations. With regards to the actual flight, the DJI Phantom 4 Multispectral (P4M) can capture all the same data as the DJI Phantom 4 Pro (P4P), therefore, just the P4M can be used to obtain the same data for this project. Flights should also be planned during periods where the plants are green to allow maximum use of the indices for classification.

During this project, there were some challenges with the Pix4D processing phases. During the Pix4D processing, index rasters should not be created using Pix4D. Instead they should be created using ArcGIS and the transparent mosaic of each band should be used to make the index calculations. It is also highly recommended to have a powerful computer processor with a lot of memory space to process these data. Extra external harddrives are needed as a lack of space caused Pix4D to crash during processing, thus losing time. Some processes that are done using ArcGIS such as hydrology and indices calculations can be automated into geoprocessing tools to save time and improve accuracy. Since this process is to be repeated often, Model Builder could save valuable time.

The distribution of vegetation seeds is determined by flow direction in the channel and on the floodplain of the meandering river (van Dijk, 2013). UAS-derived DEMs provide high-resolution datasets that accurately allow for the definition of flow direction, flow accumulation vital

parameters controlling vegetation distribution. According to Tal and Paola (2010), vegetation predominantly settles on point bars on the outer banks of rivers (floodplains), which describes the study area. The vegetation is clustered along point bars, around micro aquifers, low flood velocity areas, which also agrees with the particle distribution model for the study area. Flooding is a primary factor responsible for vegetation distribution and growth within the study area flow accumulation is a secondary factor that determines plant growth during dry and hot seasons for they accumulate and store surface water runoff. Some areas within micro aquifers appear to have no vegetation cover. These areas are highly porous and transmissive (sand and gravel) to water but lack clay material and humus that favour plant growth. They make ideal reservoirs for water storage for plants in the dry and hot seasons.

From the Random Forest Classification, of the eight species classified, giant reed clustered along the shallow water with very low energy. They are also closely associated with cattail which grows a little further into deeper water. Mesquite grows furthest away from the river banks, along with chamomile, sahara mustard, and arrow weed which all occupy the low canopy heights of the low energy flood areas where seeds tend to be deposited during floods. Cottonwood is part of the high canopy together with salt cedar. They cluster along ancient and current riverbanks while salt cedar dominates throughout the study area.Random Forest algorithm is an efficient method for vegetation classification using UAS, but some similar vegetation species such as mesquite and salt cedar still need to be verified by humans. Because even human eyes cannot describe this fluffy plant as salt cedar or mesquite. When the Random Forest model is classifying, the most important features are the DEMs, which add up to 55%. Then the most important multispectral indices are NDVI and SAVI, both accounted for eight percent.

References

Aasen, H., Burkart, A., Bolten, A., & Bareth, G. (2015). Generating 3D hyperspectral information with lightweight UAV snapshot cameras for vegetation monitoring: From camera calibration to quality assurance. ISPRS Journal of Photogrammetry and Remote Sensing, 108, 245-259. 10.1016/j.isprsjprs.2015.08.002. Retrieved from https://reader.elsevier.com/reader/sd/pii/S0924271615001938?token=413A64C294A946DFE4FEFCE8976BB48096B2EAC0EE98A539D4E2AD7B48BD2F7F15364A874A35F80745A3BD199EFB743A

Addy, S., Cooksley, S., Dodd, N., Waylen, K., Stockan, J., Byg, A., & Holstead, K. (2016) River Restoration and Biodiversity: Nature-based solutions for restoring rivers in the UK and Republic of Ireland. CREW reference: CRW2014/10, ISBN: 978-0-902701-16-8. www.crew.ac.uk/publications

Al-Najjar, H., Kalantar, B., Pradhan, B., Saeidi, V., Ueda, N., & Mansor, S. (2019). Land Cover Classification from fused DSM and UAV Images Using Convolutional Neural Networks. Remote Sensing, 11(12), Remote Sensing, Feb 2019, Vol.11(12). Retrieved from https://www.mdpi.com/2072-4292/11/12/1461

Arizona Emergency information Network. (2017, April 28). Firefighters Successful Holding Fire Lines on the Cactus Fire. Retrieved from https://ein.az.gov/emergency-information/emergency-bulletin/firefighters-successful-holding-fire-lines-cactus-fire

Arizona Emergency Information Network. (2017, April 30). Cactus Fire. Retrieved from https://ein.az.gov/keywords/cactus-fire

Baena, S., Boyd, D., & Moat, Justin. (2017). UAVs in pursuit of plant conservation - Real world experiences. Ecological Informatics. 10.1016/j.ecoinf.2017.11.001. Retrieved from https://www.sciencedirect.com/science/article/pii/S1574954117302741

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing, 114, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Centre for Agriculture and Bioscience International (CABI). (n.d.). Impacts of invasive species. Retrieved from https://www.invasive-species.org/impacts/

DJI. (2019, September 25). DJI Introduces P4 Multispectral For Precision Agriculture and Land Management. Retrieved from https://www.dji.com/newsroom/news/dji-introduces-p4-multispectral-for-precision-agriculture-and-land-management

DJI. (n.d.). Phantom 4 Pro. Retrieved from https://www.dji.com/phantom-4-pro/info#specs

Dong, P., & Chen, Q. (2018). *Lidar remote sensing and applications*. CRC Press

Esri. (2016). Hillshade function—Help | ArcGIS for Desktop. Arcgis.Com. Retrieved from http://desktop.arcgis.com/en/arcmap/10.3/manage-data/raster-and-images/hillshade-function.htm

Esri. (n.d.). An overview of the Neighborhood toolset. Retrieved from https://pro.arcgis.com/en/pro-app/tool-reference/spatial-analyst/an-overview-of-the-neighborhood-tools.htm

Fahrenkamp-Uppenbrink, J. & Funk, M. (2019). River restoration guided by research, Science :Vol. 365, Issue 6459, pp. 1260-1262. DOI: 10.1126/science.365.6459.1260-j
Flood Control District of Maricopa County. (2018, September 14). Rainfall Data. Retrieved from http://alert.fcd.maricopa.gov/alert/Rain/mean_annual_rain.png

Heintz, M. D., Hagemeier-Klose, M., & Wagner, K. (2012). Towards a Risk Governance Culture in Flood Policy—Findings from the Implementation of the "Floods Directive" in Germany. *Water*, *4*(1), 135–156. https://doi.org/10.3390/w4010135

Kattenborn, T., Eichel, J., Wiser, S., Burrows, L., Fassnacht, F.E. and Schmidtlein, S. (2020), Convolutional Neural Networks accurately predict cover fractions of plant species and communities in Unmanned Aerial Vehicle imagery. Remote Sens Ecol Conserv. https://zslpublications.onlinelibrary.wiley.com/doi/full/10.1002/rse2.146

Keeley, J. E., & Syphard, A. D. (2016). Climate Change and Future Fire Regimes: Examples from California. Geosciences, 6(3), 37. https://www.mdpi.com/2076-3263/6/3/37

Khan, N., Rastoskuev, V., Sato, Y., & Shiozawa, S. (2005). Assessment of hydrosaline land degradation by using a simple approach of remote sensing indicators. Agricultural Water Management, 77(1), 96-109. https://www-sciencedirect-com.ezproxy1.lib.asu.edu/science/article/pii/S0378377405000910

Kim-Anh, & Thanh-Hung. (2020). Soil salinity assessment by using near-infrared channel and Vegetation Soil Salinity Index derived from Landsat 8 OLI data: A case study in the Tra

Vinh Province, Mekong Delta, Vietnam. Progress in Earth and Planetary Science, 7(1), 1-16. https://progearthplanetsci.springeropen.com/articles/10.1186/s40645-019-0311-0

Lite, S. J. , Marler, R., Paradzick, C., Shorrock, D., Stromberg, J. C., White, J. M., & White., M. S. (2007). Altered Stream-Flow Regimes and Invasive Plant Species: The Tamarix Case. Global Ecology and Biogeography, 16(3), 381-393. Retrieved May 7, 2020, from www.jstor.org/stable/4139429

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. (2019). Deep learning in remote sensing applications: A meta-analysis and review. ISPRS Journal of Photogrammetry and Remote Sensing, 152, 166-177. Retrieved from https://www.sciencedirect.com/science/article/pii/S0924271619301108

Maricopa County. (n.d.). Rainfall Data. Retrieved from https://www.maricopa.gov/625/Rainfall-Data

Mesa, Arizona. (n.d.). Retrieved from https://en.wikipedia.org/wiki/Mesa,_Arizona

Mitchell, G. (2017, April 29). Latest information on the Sawmill and Cactus fires. Retrieved from https://www.azcentral.com/story/news/local/arizona/2017/04/27/what-know/307136001/

Mokhtar, E., Pradhan, S., Ghazali, B., & Shafri, A. (2018). Assessing flood inundation mapping through estimated discharge using GIS and HEC-RAS model. Arabian Journal of Geosciences, 11(21), 1-20.

National Forest Foundation. (n.d.). Lower Salt River Restoration Project. Retrieved from https://www.nationalforests.org/who-we-are/regional-offices/southernrockies/lower-salt-river-restoration-project

Nguyen, U., Glenn, E., Dang, T., & Pham, L. (2018). Mapping vegetation types in semi-arid riparian regions using random forest and object-based image approach: A case study of the Colorado River Ecosystem, Grand Canyon, Arizona. Ecological Informatics. 50. 43-50. 10.1016/j.ecoinf.2018.12.006.https://reader.elsevier.com/reader/sd/pii/S1574954118301535?token=6566F102086956B86C4D52A2138341AF66CE002652A27AFD5ADE7F482EC3B80391FE9FE54E35C7AF4657E254254FAA19

Niemeyer, J., Rottensteiner, F., & Soergel, U. (2014). Contextual classification of lidar data and building object detection in urban areas. ISPRS Journal of Photogrammetry and Remote Sensing, 87, 152–165. https://doi.org/10.1016/j.isprsjprs.2013.11.001

Northern Arizona Invasive Plants. (n.d.). Tamarisk. Retrieved from https://www.nazinvasiveplants.org/tamarisk

Orengo, H. A., & Petrie, C. A. (2018). Multi-scale relief model (MSRM): a new algorithm for the visualization of subtle topographic change of variable size in digital elevation models. *Earth Surface Processes and Landforms*, *43*(6), 1361–1369. https://doi.org/10.1002/esp.4317

Palmer, M., & Ruhi, A. (2019). Linkages between flow regime, biota, and ecosystem processes: Implications for river restoration. Science, Vol.365(6459).ISSN: 00368075, E-ISSN: 10959203, DOI: 10.1126/science.aaw2087

Pix4D SA,(2017) Pix4Dmapper 4.1 User Manual. Retrieved from https://support.pix4d.com/hc/en-us/sections/360003718992-Manual

Robles, M. D., Turner, D. S., & Haney, J. A. (2017). A century of changing flows: Forest management changed flow magnitudes and warming advanced the timing of flow in a southwestern US river. PloS one, 12(11), e0187875. https://doi.org/10.1371/journal.pone.0187875

Santos, W., Silva, B., Oliveira, G., Volpato, M., Lima, J., Curi, N., & Marques, J. (2014). Soil moisture in the root zone and its relation to plant vigor assessed by remote sensing at management scale. Geoderma, 221-222, 91-95. https://www-sciencedirect-com.ezproxy1.lib.asu.edu/science/article/pii/S0016706114000159

SEINet Arizona - New Mexico Chapter. (n.d.). Retrieved from http://swbiodiversity.org/seinet/taxa/index.php?taxon=344

USDA (2014). Field Guide for Managing Saltcedar in the Southwest. Retrieved from https://riversedgewest.org/sites/default/files/resource-center-documents/Field_Guide_for_Managing_Tam_SW.pdf

USDA. (n.d.). History and Development. Retrieved from https://www.fs.usda.gov/detail/tonto/about-forest/?cid=fsbdev3_018924

Wahab, I., Hall, O., & Jirstrom, M. (2018, August 16). Remote Sensing of Yields: Application of UAV Imagery-Derived NDVI for Estimating Maize Vigor and Yields in Complex Farming Systems in Sub-Saharan Africa. Retrieved from https://www.mdpi.com/2504-446X/2/3/28/pdf-vor

Ward, J., Tockner, K., & Schiemer, F. (1999).Biodiversity of floodplain river ecosystems: ecotones and connectivity. Regulated Rivers: Research & Management, 15, 125-139