LOCATING TURF AND WATER FEATURES IN THE LAS VEGAS VALLEY, NEVADA, USING REMOTE SENSING TECHNIQUES AND GIS

Judy Brandt
GIS Analyst
Southern Nevada Water Authority
Groundwater Resources Department
Data Resources Division
100 City Parkway, Suite 700
Las Vegas, Nevada 89106
judy.brandt@snwa.com

ABSTRACT

The Southern Nevada Water Authority (SNWA) was established in 1991 by the State of Nevada to secure future water resources of Southern Nevada communities, including the fast-growing city of Las Vegas. Part of this effort involves conservation of water in this dry desert community. Nearly 70% of water consumed is due to outdoor use—primarily irrigation of turf (grass). Additionally, some water is lost to evaporation from surface sources, including man-made water features such as swimming pools.

SNWA has established programs which offer incentives to water consumers to reduce turf area in landscaping and to use swimming pool covers. In order to market conservation programs to the appropriate consumers, object-based and pixel-based image classification techniques were used to identify turf and swimming pools in the Las Vegas Valley. High resolution (6 inch), 3-band color infrared digital aerial imagery was collected in June of 2006 over the Las Vegas metropolitan area and surrounding communities. Using ERDAS Imagine™, a single band "diversity layer" was produced from each images' DN values and this layer was stacked into the three band CIR image. Visual Learning Systems' Feature Analyst™ extension for ESRI ArcGIS™ (9.1) was used to classify vegetation into trees, turf and shadows. An overall accuracy of 92% confirmed that SNWA could have confidence in the final dataset. A normalized band difference was used to determine the location and area of swimming pools and man-made ponds and lakes in the Las Vegas Valley. A GIS and "decision-based" geo-processing produced a "clean", final swimming pool dataset.

GIS and common geo-processing tools were used to examine the relationships between vegetation layers, swimming pool data and existing municipal data. This allowed for identification of high-density turf neighborhoods and properties with swimming pools, and enhanced target-marketing for SNWA's turf-reduction and pool cover programs.

PURPOSE

The purpose of identifying turf areas is directly related to the progress of SNWA's water conservation programs. In order to encourage the replacement of turf with more drought-tolerant and less water-consumptive landscape, the SNWA established the Water-Smart Landscape (WSL) program in 2002. The WSL program provides a cash incentive for commercial and residential customers to replace turf with approved, drought-tolerant landscaping (Figure1). The total amount of turf converted peaked in 2004 but slowly declined thereafter until the cash incentives were recently increased. Currently, there is a $1.5 per square foot cash rebate for turf replaced. The program is successful, with tens of millions of square feet of turf converted and billions of gallons of water saved annually. SNWA would like to refine these programs to make them more cost efficient and effective in reaching the right demographic. The SNWA Conservation Division and Data Resources Division developed an approach to this problem. To determine how well the WSL and other lawn conversion programs are working, obtaining a baseline value of how much turf is in the valley is necessary. By locating exactly where the turf is, target marketing can be used to push conservation programs in areas where larger concentrations of turf are located. SNWA determined the most efficient way of finding turf in the LV Valley was to obtain high resolution multi-spectral imagery and perform analysis that would allow wholesale identification of the areas of turf and other vegetation.

In addition to locating turf, identifying swimming pools was also a priority of SNWA's water conservation efforts. In terms of water conservation, swimming pools rank high as a water waster. Though the total area of pools is not nearly that of turf, there are concerns about the loss of water through evaporation from the open water surface.

SNWA has instituted programs where rebates are given for the cost of pool covers. Permits are issued for pools by Clark County, but those records are not recorded in a timely way and they are not required for above-ground pools. It was expedient to use the 2006 high resolution imagery to obtain current data on swimming pools in the Las Vegas Valley and possibly use the data for target marketing of the pool cover program.

Figure 1. Photo at left shows residential landscape before WSL program enrollment. The photo at right shows the same residence after conversion of turf to desert landscaping. *Photos: Courtesy SNWA Conservation Division*

IMAGERY

The high resolution (6") digital imagery collected by Digital Mapping, Inc. (DMI), for SNWA comprised 3 bands in the green (500-650nm), red (590-675nm) and near-infrared wavelength (675-850nm) ranges (Figure 2). The original data was collected as a 12-bit product using a Digital Mapping Camera System (equipped with Airborne GPS/IMU) by Intergraph-Z/I Imaging. The study area consisted of the metropolitan Las Vegas area including the cities of Boulder City, Henderson and North Las Vegas and unincorporated Clark County, including the Las Vegas Strip. The original 12-bit images were mosaicked, color-balanced, ortho-rectified and tiled by DMI. A 12-bit, compressed product supplied to SNWA was tiled to match the Public Land Survey Sections of Clark County, Nevada. The image tile name corresponds to the book number (township and range) and section number, which was unique to each tile. Each tile covers approximately one square mile. The study area for the turf project consisted of approximately 475 tiles (Figure 3). The digital data was collected in June of 2006, when vegetation was in the peak of health.

Figure 2. Typical CIR Image tile, *DMI, June, 2006.*

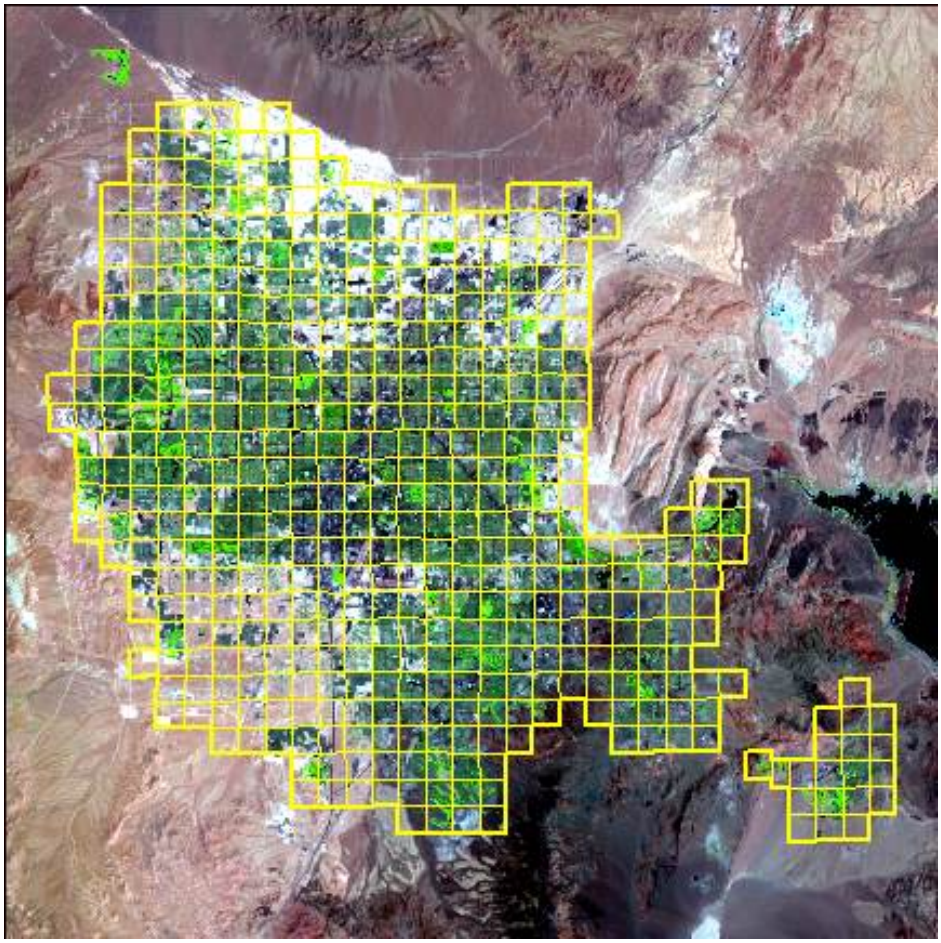

Figure 3. Landsat 5 TM (June, 2004, Bands 5-4-3) image showing the Las Vegas Valley. The study area is outlined in heavy yellow and image tiles in yellow.

METHODOLOGY - VEGETATION ANALYSIS

Part of the challenge of this project involved handling, processing and storage of the large quantity of original data and processed data. The image processing was done on a Sun PC Dual Core AMD Opteron™ with an MS Windows Professional x64 operating system. In ERDAS Imagine™, the 12-bit images were rescaled to an uncompressed, unsigned 8-bit .tif dataset compatible with ArcGIS™ v9.1. Approximately 3 months were spent in determining the optimum methodology for analyzing the imagery, executing the methodology and extracting the data desired.

Image Collection

The digital image data was collected over a period of two weeks, and was flown at different altitudes due to flying restrictions in some parts of the study area. This, and other factors such as the urban-suburban nature of the terrain, the type of soil in the LV Valley and the amount of vacant land caused some tiles to have a high degree of color and texture variation within them (Figure 4) in spite of DMI's efforts at color-balancing. From tile to tile the variation could also be quite dramatic. This extreme variation made it necessary to produce a mosaic image suitable for creating a classification training set that would lead to an acceptable supervised classification. Six adjacent image tiles, oriented perpendicular to the flight lines, were mosaicked using Imagine.

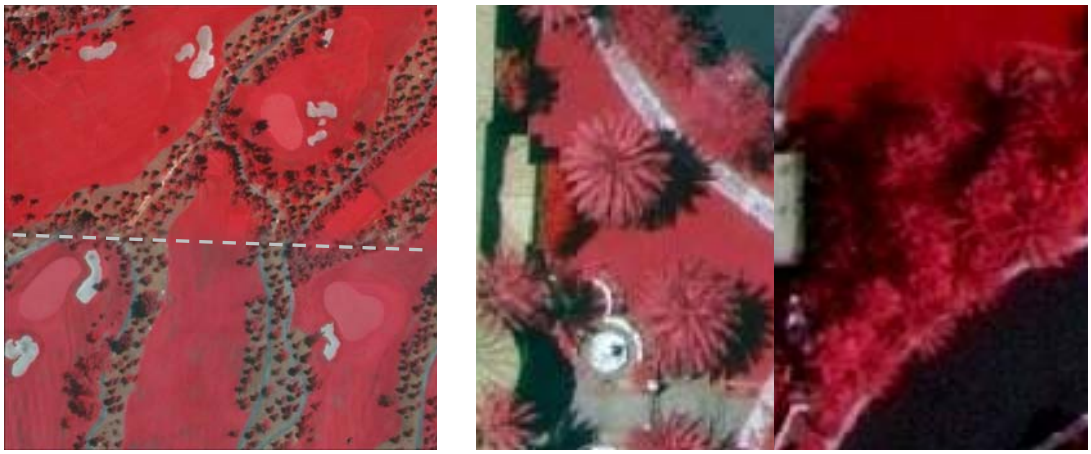

Figure 4. The image tile on the left is an example of color variation within a single tile. Here the image is a mosaic of data collected on two different days or times of day. The image tiles on the right, shown at the same scale, illustrate color and texture differences from image to image. *DMI, June, 2006*

Classification Model

The goal of the classification was to identify three classes of cover: trees (including shrubs), turf and vegetation in shadows. These three were chosen because the goal of the project was to extract turf from the image, including turf in shadow areas (between buildings, in tree shadows), and to learn about the relationship with or association with other vegetation. Since the imagery was of high resolution, no field work was needed to determine training areas. From examination of the imagery it was clear that some types of trees appeared very similar to some areas of turf, and some types turf were very different from each other. To the extent possible, when choosing the training areas, the entire range of color and texture in each class was used. Because of the quantity of the images, the time required for processing and project deadlines, there was not time to "clean up" the classification for individual image tiles, combining multiple classes, etc. The approach was to come up with the best training set and model, run the classification on every one of the 475 images, perform an accuracy assessment and present the results before the deadline.

The Visual Learning Systems' (VLS) Feature Analyst™ extension for ArcGIS™ was used in image classification. This software was selected because of its ability to extract the turf, tree and shadow classes efficiently. In the urban landscape, the vegetated areas can be seen as "objects", which the Feature Analyst™ software was able to extract relatively well in a one-step process. Using the mosaic image mentioned previously, 25-30 training polygons were identified for each class and were merged into a single training shapefile.

The Feature Analyst™ classification model built by the user allows for input of a number of settings and parameters. Included in this are feature type, input bands, input representation, and "learning options". Many classifications with different patterns, feature models and pixel aggregates were tried on the mosaic and on samples of the images. The total processing time was noted and the classification results were compared to the imagery. It was this "trial and error" process that determined the final model for the analysis. The most successful model used was for the extraction of a "natural" feature type (for example, trees). The search pattern (input representation) for detection of objects was a "bulls-eye 3", 13 pixels wide. A "learning option" that strongly influenced the classification was the aggregate size. The aggregate size used in the final model was 225 pixels, which means that groups of "qualified" pixels smaller than this were not classified. Only the qualified pixel groupings were classified and appear in the resultant dataset

The trial classifications revealed a need to more clearly define the boundaries of features and perhaps add more textural information, as well. A "diversity" neighborhood (5x5 matrix) function was performed on the mosaic image using ERDAS Imagine 9.0. The resulting 3-band "diversity" image was then summed, and the resultant single band image (Figure 5) was stacked onto the 8-bit three band image. This "final" 4-band mosaic image file (.img) was used to create a new training set and was then classified in ArcGIS™ using the preferred Feature Analyst™ model. This produced the best classification result. Because of this success, all of the images were batch-processed to add the "diversity" layer and the batch processor in Feature Analyst™ was used to classify the resultant 4-band images using the preferred model and the mosaic classification results as a "training set".

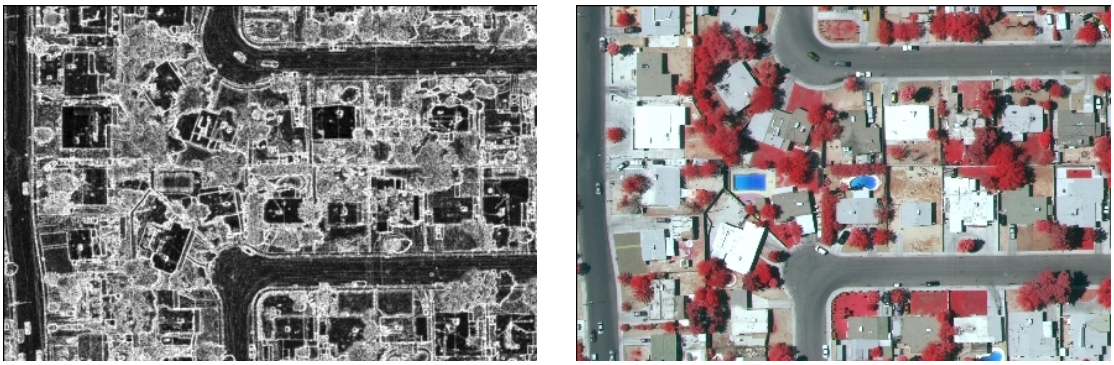

Figure 5. The image at left shows the diversity layer, the image at right shows the same area in the original 3-band image. *DMI, June, 2006*

The dataset resulting from the Feature Analyst™ classification process was a single shapefile containing each of the separate class polygons. Each class was identified by a specific attribute and could be symbolized in ArcMap as such. After the classification, the 475 shapefiles needed some work to determine which shadows were likely turf (in shadows), which were shadows in tree canopies and which would have no association with vegetation. The data also needed to be related with other GIS data that would be useful in determining which housing developments had more turf per residence, or which homeowners had copious amounts of turf and vegetation in their yards.

GIS Analysis

Using the Modeler in ArcToolBox a single geo-processing model was constructed that would accomplish the necessary analysis. The model was then exported to a Python script to enable looping through the 475 tile shapefiles. Unfortunately, memory leaks were common to ArcGIS™ 9.1 when geo-processing with Python, and this frequently caused the script to after processing only a few datasets. With the release of ArcGIS™ version 9.2, the memory leaks were no longer problematic and the script ran smoothly through all of the datasets. The scripts had to be run on a 32bit Intel Pentium processor machine with MS Windows XP Professional™. A 64-bit dual core machine was available, but the Python script with the ArcGIS™ 9.2 geo-processing tools did not work in a 64-bit OS. The model contained all of the necessary steps required to finalize the data. It integrated the vegetation data with the most current Clark County Assessor's Office parcel data and distributed shadow polygons into classes of tree, turf or "neither" (Figure 6). The criteria for determining whether shadows belonged to the tree or turf class were as follows: If a shadow polygon was within 6 pixels of a turf polygon it was considered turf. Of the remaining shadow polygons, if the polygon was within one pixel of a tree polygon, it was then considered a tree. Any shadow polygons not meeting the criteria were considered non-vegetation. The ArcGIS™ "Select by location" and "Selection" geo-processing tools were used to accomplish this task within the script.

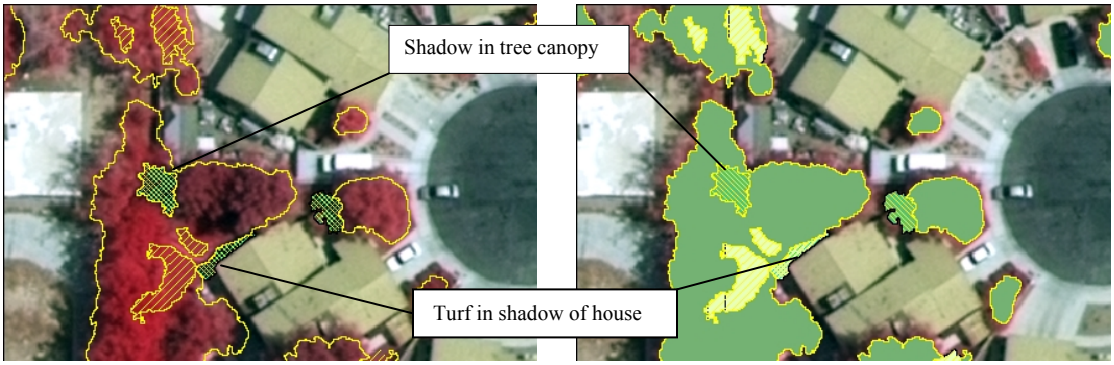

Figure 6. The figure on the left shows the original classification data, with grass shown as hatch, trees as a yellow outline and shadow as a double hatch. On the right the grouping of the shadows into respective turf and tree classes is complete, with trees shown in green and turf in yellow.

With the vegetation data in a final form, the next step was to give it more meaning by combining it with existing GIS data. The five digit combination of the book and section number of the public land survey is an attribute in the Clark County Assessor's office parcel layer and was, by design, part of the name of the image and resulting image classification shapefile. This facilitated a selection of the parcel layer data corresponding to the image data area and enabled the use of the ArcGIS™ intersect tool to combine parcel layer with the image class data (Figure 7). After the intersection, the vegetation class polygons had all of the attributes of the parcel data as well as the vegetation-type information. The parcel number is important for linking the vegetation data to many types of municipal data and enables the SNWA to conduct a market analysis for the Water Smart Landscape turf reduction program.

Geo-processing time was 15-20 minutes per dataset because the large number of polygons produced from the analysis and because each polygon had a large number of vertices. Edges of polygons were simplified (vertices removed) to facilitate some of the processes. However, the final product had all of the original vertices. The final datasets for each image tile included one tree shapefile, one turf shapefile, one shadow shapefile, one shadow shapefile that would be merged with the trees, and one shadow shapefile that would be merged with the turf. The shadow datasets were manually merged with the respective turf and tree datasets using the ArcGIS™ "Merge" tool in batch mode. The 475 turf and tree section tile datasets were then merged by book number and imported into personal geodatabases as polygon feature classes. Each polygon feature had a parcel number as an attribute to enable joining with other municipal data containing property owner information, land-use information and water consumption data.

Figure 7. This image is overlain by turf polygons in yellow hatch, tree polygons outlined in yellow and parcels outlined in blue. Some parcel boundaries are in the front yards, showing how the imagery did not perfectly line up with Clark County Geographic Information System Management Office GIS layers.

ACCURACY ASSESSMENT

While examining the classification results, it became clear there were some errors in the classification (Figure 8). In order to determine the extent of any errors and the accuracy of the data, an accuracy assessment was performed. Using ArcGIS™, random points were generated over the entire study area in each of three cover types: turf, tree and non-vegetation. The non-vegetation area essentially consisted of everything within the study area (see Figure 3) that was not previously classified as tree or turf. The pixel at each point location was examined using the imagery available. The high resolution imagery enabled positive visual identification of over 95% of the points. A very small percentage of the points had to be identified in the field. Field identification was somewhat biased because of accessibility issues. Only front yards of private properties and public property could be accessed well enough for a firm identification in the field. An error matrix was generated as the accuracy data was accumulated.

As shown in Table 1, a total of 1417 point locations were examined. The overall accuracy was calculated to be 91.39%, and the K_{hat} Kappa analysis (Congalton et al., 1999) results calculated at 86.92%. Based on the producer and user accuracy shown, the overall accuracy and the Kappa analysis results, this data was given a very acceptable level of confidence, enabling its use in decision-making and in marketing WSL programs.

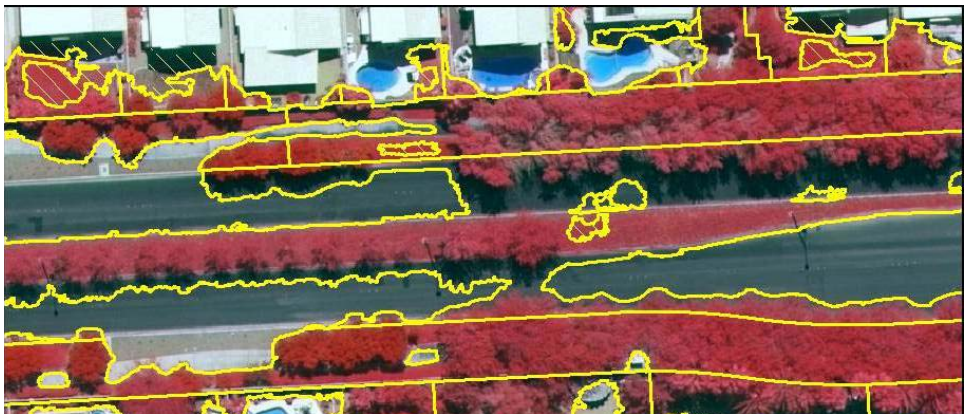

Figure 8. The tree polygons are outlined in yellow along the roadway. Some of the polygons extend beyond the tree canopy and cross the roadway to the canopy on the other side.

Table 1. Error/Accuracy matrix, including overall accuracy and K_{hat} results

	Turf	**Tree**	**Non-Vegetation**	**Row Total**	**User Accuracy**
Turf	321	49	1	371	**86.52%**
Tree	28	489	9	526	**92.97%**
Non-Vegetation	10	25	485	520	**93.27%**
Column Total	359	563	495	1417	
Producer Accuracy	**89.42%**	**86.86%**	**97.98%**		
Overall Accuracy = 91.39%					
K_{hat} = 86.92%					

DATA APPLICATION – TURF EXTRACTION

The accuracy assessment confirmed to SNWA that the data was meaningful and usable for market analysis. The ability to join to parcel data enables the determination of which parcels had the densest cover of turf or vegetation and provides address information on the parcel owners. After completion of this analysis the SNWA Conservation Division executed a targeted direct mailing promoting the WSL program to single-family households in the Las Vegas Valley. The top 30,000 parcels, in terms of turf area, were targeted. The response to the mailing was more than double that of a comparable, non-targeted mailing previously done by SNWA.

Using the GIS data to identify turf in medians and rights-of-way, SNWA can target municipalities that may still be maintaining turf in these areas. The data can be linked to consumption data, which would enable an examination of the relationship between landscaping type and total water consumption for a parcel. This could help confirm whether the focus of the conservation efforts is where it should be.

METHODOLOGY – SWIMMING POOL EXTRACTION

Band Analysis

The extraction of swimming pools from the 2006 aerial data was a much different process from that of vegetation. In an earlier, unpublished study of high-resolution satellite data, it was found, accidentally, that the swimming pools could be extracted using comparative band values. In the study of the 2006 aerial imagery, a normalized band difference was used in the analysis. Because of the difference in the spectral absorption of the NIR and the green band, it seemed logical that the difference of those two band values could help determine the area and boundaries of both natural and man-made water bodies. The images were processed to come up with an image (.img) containing values determined by the following formula:

$$\frac{\text{NIR} - \text{GREEN}}{\text{NIR} + \text{GREEN}}$$

The values were normalized because it was thought that with the DN variability across the 475 tiles could lead to erroneous results. Normalization should lead to a more universal result that could be applied to all the tiles. Using ERDAS Imagine™ a model was built using the above formula, and all 475 of the images were batch processed. The resultant image (.img) files were then fed into an ArcGIS™ geo-processing model which refined the swimming pool extraction.

GIS Analysis

Through trial and error, a simple "decision tree" was developed and incorporated into the ArcGIS™ geo-processing model and that model was converted to a Python script and run as a batch process. The process took into account the pixel values of the image files. Within the model, the images were reclassified twice, to produce two separate datasets. One dataset encompassed a slightly broader range of values than the other:

Value Range Dataset A: -1.0 - -0.08
Value Range Dataset B: -1.0 - -0.10

Each dataset was converted to polygon feature classes within a geodatabase. Examination of this data revealed an unacceptable number of reflections off of features such as parking lot roofs and car windshields classified incorrectly as pools. The vast majority of these polygons were filtered out of the datasets using a query based on the total area and the ratio of the area to the polygon circumference. The actual values were determined from those which yielded the most accurate results based on trial and error on a sample of the data. The units were in square feet:

Filter for Dataset A: Shape_Area > Shape_Length AND Shape_Area > 7
Filter for Dataset B: Shape_Area > Shape_Length AND Shape_Area > 9

Through our analysis of the data, it was found that the pools had a cluster of pixels at their core which fell into the dataset B value range, but the total surface area of the pools encompassed pixels which fell into the dataset A value range. The polygons falling within the narrow value range (B) determined which polygons with the wider value range (A) would appear in the final dataset. If a polygon in dataset A did not touch a polygon from dataset B, it was determined that it was not a swimming pool. The ArcGIS™ "Select By Location" tool was used to perform this process. Essentially, only "qualified" polygons from dataset A were included in the final data.

Integration With Municipal Data

The final step of the geo-processing combined the Clark County Assessor's Office parcels with the swimming pools through performance of an "Intersection" of the two polygon datasets. In this step, the polygon areas from each dataset were merged into one where they overlapped, and the feature attributes from each dataset were retained. Parcel numbers were then attributed to each pool, which enabled joining with other municipal datasets.

Manual corrections of the entire final swimming pools dataset showed few errors, though no formal assessment was done to determine the accuracy of the data. Most of the errors included true water bodies such as golf course ponds. These were minimal and easily filtered out to produce a "pure" swimming pool dataset either manually or

through a query of land use. The data was used by SNWA to establish a base from which to begin study of trends in number and size of pools in the Las Vegas Valley. Also, comparing the Clark County Assessor's pool permit data to this data gave an insight into how complete and accurate (in terms of surface area) the County data is and allowed SNWA analysts to consider the contribution of above ground pools (not permitted) to the surface water evaporation in the Las Vegas Valley.

CONCLUSIONS AND FUTURE WORK

Conclusions

The object-based imagery analysis, using the VLS Feature Analyst™ extension for ArcGIS, proved to be a very acceptable methodology for extracting vegetation from high-resolution digital aerial imagery. Trial and error helped refine the process and define the classification model. The addition of a fourth, new layer into the original, color-balanced and tiled 3-band image data yielded more accurate results than the analysis of the original imagery, which was nearly acceptable on its own. GIS analysis distributed shadows into the proper class of vegetation (or non-vegetation) and integrated the data with existing municipal GIS datasets. The integration of datasets and the accuracy of the vegetation data gave it meaning and viability. This enabled improved marketing of Water Smart Landscape turf reduction programs and created an understanding of the distribution of vegetation in the Las Vegas Valley.

A basic normalized band ratio using the green and IR bands of the digital aerial 3-band imagery provided the basis for extracting swimming pools and other water features. GIS analysis using ArcGIS™ tools refined the original data and helped successfully develop a "clean" GIS swimming pool dataset. This data supplemented existing, but not entirely complete or accurate permit data available from the Clark County Assessor's Office. SNWA analysts now have the data necessary to determine how much of a contribution swimming pools make to water loss in the Las Vegas Valley and can identify trends in pool size and geographic distribution. They may also use the data to successfully market pool cover programs, hopefully reducing water loss through evaporation.

Future Work

SNWA has obtained aerial imagery of the Las Vegas Valley from May 2007 and 2008 and is performing analysis to extract both vegetation and swimming pools, as before. This was deemed necessary, in order to determine if the pace of turf reduction is comparable to the planting of new turf due to construction of new homes and landscapes. Clearly the goal of SNWA is to reduce the amount of total turf in the Las Vegas Valley. Continued monitoring of turf can only enhance its efforts and give it a meaningful measure of how closely it is approaching that goal. At the same time, monitoring pools can tell SNWA whether pools are increasing or decreasing in size and frequency, and how patterns of pool use may be changing as landscaping patterns change.

REFERENCES

Congalton, R.G. and K. Green, 1999. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices.* Lewis Publishers, Boca Raton, FL, pp50.