Mapping Impervious Surfaces Using Object-oriented Classification in a Semiarid Urban Region

Zachary P. Sugg, Tobias Finke, David C. Goodrich, M. Susan Moran, and Stephen R. Yool

Abstract

Mapping the expansion of impervious surfaces in urbanizing areas is important for monitoring and understanding the hydrologic impacts of land development. The most common approach using spectral vegetation indices, however, is difficult in arid and semiarid environments where vegetation is sparse and often senescent. In this study object-oriented classification of high-resolution imagery was used to develop a cost-effective, semi-automated approach for mapping impervious surfaces in Sierra Vista, Arizona for an individual neighborhood and the larger sub-watershed. Results from the neighborhood-scale analysis show that object-oriented classification of QuickBird imagery produced repeatable results with good accuracy. Applying the approach to a 1,179 km² region produced maps of impervious surfaces with a mean overall accuracy of 88.1 percent. This study demonstrates the value of employing object-oriented classification of high-resolution imagery to operationally monitor urban growth in arid lands at different spatial scales in order to fill knowledge gaps critical to effective watershed management.

Introduction and Background

Recent trends of population in-migration related to environmental amenities in Arizona and many other parts of the Rocky Mountain region of the US have been associated with high rates of urbanization and land development (Vias and Carruthers, 2005). Impervious surfaces (materials that prevent the infiltration of water into soil (Slonecker *et al.*, 2001)) are created by construction activities, affecting land surface temperature, water quality, and watershed properties directly. Increases in the amount and distribution of impervious surfaces in rapidly urbanizing areas can produce potentially significant changes in hydrological processes in watersheds by altering runoff regimes, increasing peak flows, and degrading water resources (Arnold, Jr. and Gibbons, 1996; Kennedy *et al.*, 2013; Shuster *et al.*, 2005). Additionally, the spatial distribution of impervious areas is an important descriptor of the physical content of urban environments (Chormanski *et al.*, 2008; Shuster *et al.*, 2005). Mapping impervious surfaces with remote sensing techniques is an effective way to quantify impervious cover (Slonecker *et al.*, 2001; Weng, 2007) and thereby improve understanding of the impacts of urbanization on runoff processes. The most common approach using spectral vegetation indices, however, is problematic in arid and semiarid environments where vegetation is patchy and often senescent.

This paper describes a method for mapping impervious surfaces using supervised object-oriented classification of high-resolution imagery for an urbanizing semi-arid area. Insights are provided at the scale of an individual neighborhood as well as the larger sub-watershed to show that despite utilizing high-resolution imagery, the method is not limited to only small geographical areas. The first section provides background on the use of object-oriented classification approaches for detecting impervious surfaces and identifies the need for applications to arid and semi-arid locations, followed by a description of the study areas and imagery used. The next section describes the methods and results from the neighborhood scale classification (phase 1); then, the methods, results, and errors and limitations of the regional scale classification (phase 2) follow. The final section offers conclusions and recommendations for refining the classification method.

Object-oriented Approaches to Mapping Impervious Surfaces

Earlier strategies for mapping impervious surfaces are based largely on user-guided, manual delineation (Lee and Heaney, 2003; Shuster *et al.*, 2005). The advantage of this method is its ability to distinguish between directly and indirectly connected impervious areas, which is important information for hydrologic modeling. The major disadvantage, however, is the time and effort required to produce delineations, thus limiting application to small areas (McMahon, 2007). A secondary drawback is that the digitization of impervious areas by hand can affect data consistency and accuracy.

Recent remote sensing approaches for automated mapping of urban impervious areas frequently use spectral vegetation indices as proxies for imperviousness, assuming for example that vegetated areas represent pervious surfaces (Bauer *et al.*, 2002; Sawaya *et al.*, 2003; Thanapura *et al.*, 2007). Proxies are thus based on indices such as Normalized Difference Vegetation Index (NDVI), where:

$$NDVI = (DN_{NIR} - DN_{RED}) / (DN_{NIR} + DN_{RED}) \quad (1)$$

Zachary P. Sugg is with the School of Geography and Development, University of Arizona, 443 Harvill Building, University of Arizona, Tucson, AZ 85721 (zsugg@email.arizona.edu).

Tobias Finke is a Consultant; 756 E. Winchester Street, Ste. 400, Salt Lake City, UT, and formerly with School of Geography and Development, University of Arizona, Tucson, AZ 85721.

David C. Goodrich and M. Susan Moran are with the USDA ARS Southwest Watershed Research Center, 2000 E. Allen Road, Tucson, AZ, 85719.

Stephen R. Yool is with the School of Geography and Development, 435C Harvill Building, University of Arizona, Tucson, AZ 85721.

Photogrammetric Engineering & Remote Sensing
Vol. 80, No. 4, April 2014, pp. 343–352.
0099-1112/14/8004–343

doi: 10.14358/PERS.80.4.343

and DN_{NIR} = digital number of the near infrared (NIR) band, DN_{RED} = digital number of the red spectral band, and NDVI is sensitive to vegetation cover. Approaches solely based on spectral vegetation indices and assumptions about the relation between vegetation and imperviousness, however, are not fully suitable for arid and semi-arid urban areas. This is because vegetation is sparse and often senescent and xeriscaping is becoming more popular (Colby and Jacobs, 2007). This lack of suitability is due in part to an important limitation of NDVI: its sensitivity to soil background in sparsely vegetated areas (Huete, 1988). Pure spectral classification techniques appear least successful (50 percent) and spectral classification followed by contextual modeling provides higher accuracies (80 percent), but demands significant preparation time (Thomas *et al.*, 2003).

An alternative approach for detecting impervious surfaces is object-oriented analysis, which utilizes both the spectral characteristics of pixels and their spatial arrangement and context within an image (Weng, 2012). Although specific algorithms vary and many are proprietary, object-oriented image analysis systems share at least two common mechanical bases: (a) image segmentation, i.e., the grouping of pixels into recognizable features, and (b) rule-based classifiers that assign meaningful labels to categories of features based on a variety of characteristics (Lang, 2008). This approach leads to a clear advantage of object-oriented classifiers for imagery with small ground sample distances (GSD) of one meter or less (Thomas *et al.*, 2003). While a 15 m GSD represents a distribution of trees as one class, a 60 cm GSD produces 625 spectral samples of this class in the same area, including tree structures, shadows, and soil. The much smaller GSD thus tends to yield high intra-class variances, suggesting traditional per-pixel classifiers are not well suited for such images (Kressler *et al.*, 2001; Thomas *et al.*, 2003). Because a single pixel with small GSD represents a subset of a logical class and not several different classes, object-oriented classification includes information from surrounding pixels in the pattern recognition process, increasing achievable accuracies from high-resolution images (Thomas *et al.*, 2003).

Unlike traditional pixel-based classifiers, these object-oriented systems also allow for the incorporation of expert knowledge (Lang, 2008; Platt and Rapoza, 2008), e.g., through parameterization of classification rules based on the expertise of the analyst. A review of the multitude of applications of object-oriented analysis systems is not possible here, but see Blaschke *et al.* (2008) for a diverse compilation. Overall, studies have shown that the relative advantage of object-oriented classification is its ability to produce acceptable accuracies (70 percent) with a relatively low amount of analyst input (Thomas *et al.*, 2003).

Overwatch Systems LTD's Feature Analyst®[1] (FA) is a proprietary object-oriented classification software package that has been used to detect impervious surfaces with high accuracy using very high-resolution orthoimages (Miller et al., 2007; Miller et al., 2009) and QuickBird imagery (Tsai et al., 2011). While these recent studies demonstrate the high accuracies possible, they were all conducted in humid locales. It has not yet been demonstrated what kind of accuracies may be achieved in arid or semi-arid settings with much sparser vegetation cover using the combination of object-oriented classifiers and imagery with small GSD.

[1] Use of a company or product name does not imply approval or recommendation of the product to the exclusion of other products which may also be suitable.

Present Study: Applying Object-oriented Classification to a Semi-arid Location

Given the recent advances in object-oriented classification algorithms and the critical need to understand the eco-hydrologic impacts of increasing imperviousness in urban environments, this study used FA to develop a method for classifying impervious surfaces in the rapidly developing town of Sierra Vista in semiarid southeastern Arizona using QuickBird imagery from 2007 and 2009. Sierra Vista is an ideal site for testing this method because of its location in a semiarid region. The physical complexity of urban areas makes high-resolution images especially useful because detailed urban structures can be resolved (Kressler *et al.*, 2001; Thomas *et al.*, 2003). The objectives of the study were to (a) develop semi-automated methods using imagery with small GSD and an object-oriented classification for extracting impervious areas in arid environments from satellite images for a single subdivision; (b) adapt those methods to classify impervious surfaces to the larger sub-watershed; and (c) produce a high-resolution map of impervious surfaces for the Sierra Vista sub-watershed suitable for input to hydrologic models. The methods developed in this study could be used periodically (in conjunction with the releases of census data, for instance) to monitor the urbanization of the area and other rapidly growing arid lands in tandem with available hydrological data to fill knowledge gaps critical to the effective management of watersheds and riparian zones. For example, it will lead to a better understanding of how continued urban growth may affect storm water runoff and its utilization for groundwater recharge.

Study Sites and Imagery

Study Areas

In order to develop the classification approach and also determine its effectiveness when applied to a larger region, this study proceeded in two phases at two geographic scales. The first phase was at the neighborhood level where the focus was on the La Terraza subdivision, a 13 ha gated community with homes on lots ranging from 1,600 to 2,600 m^2. Homeowner association regulations mandate low water landscaping, producing a homogenous xeriscaped yard design with little vegetation. Gravel mulch covers most of the pervious areas (Figure 1) typical of neighborhoods in municipal Sierra Vista.

The second phase of the study concentrated on a larger spatial scale consisting of a 1,179 km^2 section of the Upper San Pedro Watershed, containing the developed areas of Sierra Vista and the Ft. Huachuca military installation (Figure 2). This study area is bounded to the south by the US - Mexico border and to the west, north, and east by the Sierra Vista sub-watershed boundary. The City of Sierra Vista has experienced unprecedented urbanization resulting in increasing land area developed for residential housing. The Upper San Pedro Watershed contains the last free-flowing desert river system in the United States and parts of the watershed are protected within the San Pedro Riparian National Conservation Area, which supports a large number of bird species and related nature-based tourism activities (Goodrich *et al.*, 2000).

Imagery and Preprocessing

The local and regional scale analyses were based upon one QuickBird image acquired on 23 February 2007 and six images from 26 December 2009, respectively. The 2007 imagery was chosen because it was the latest available at the time that imaged the La Terraza development. The December 2009 acquisition was timed to coincide closely with the 2010 census. All images include a panchromatic channel (445 - 900 nm) with a nominal GSD of 0.61 m and four multispectral channels (blue: 450 - 520 nm, green: 520 - 600 nm, red: 630 - 690,

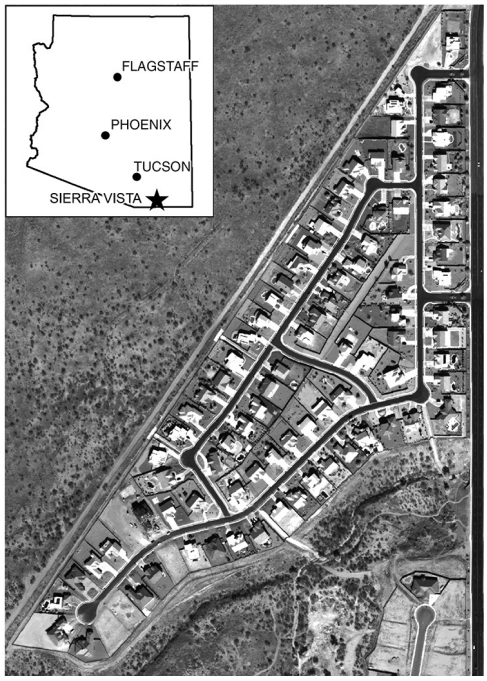

Figure 1. QuickBird image of La Terraza neighborhood in Sierra Vista, Arizona. Single-home lots exhibit xeriscape landscaping characterized by gravel cover and drought-tolerant plants that are often senescent. Inset adapted from Kennedy *et al*. (2013).

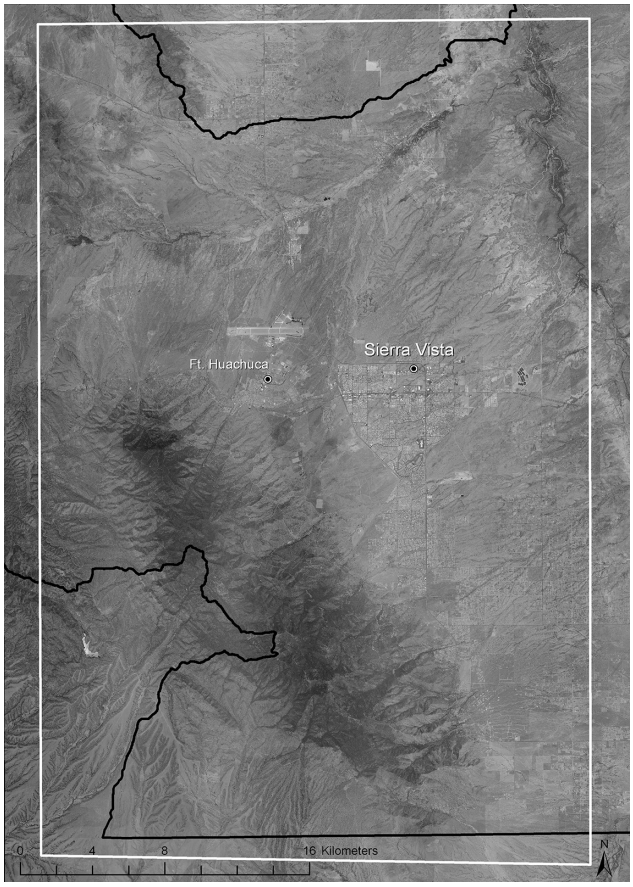

Figure 2. Sierra Vista sub-watershed boundary shown in black. Area covered by the six 2009 QuickBird images outlined in white.

NIR: 760 - 900 nm) with a GSD of 2.44 m. The 2007 image was rectified to sub-pixel accuracy, while the much larger 2009 scenes were rectified to an accuracy below 1.5 pixels. For both the 2007 and 2009 images the multispectral bands were pan-sharpened using the high-pass filter sharpening algorithm in ERDAS Imagine® to improve their spatial resolution from 2.4 m to 0.6 m. This ensured that all subsequently derived data sets were of the same resolution and provided better pattern information for the object-oriented classifier. The 2007 image was cropped to the boundary of the La Terraza neighborhood. The 2009 images were cropped to the boundary of the Sierra Vista sub-watershed.

Phase 1: Neighborhood-scale Impervious Surface Classification for an Arid Locale

Feature Extraction Using Object-oriented Classification Software

Object-oriented classification software is available from several vendors. As with other object-oriented image classification products, FA automates the extraction of features from remotely sensed imagery in order to overcome the problems associated with manual delineation, i.e., that it is laborious and time consuming, which can impose high labor costs and varying levels of accuracy (Blundell and Opitz, 2006). FA is an extension within commercial GIS software with a user interface in which "[t]he user gives the system (computer program) a sample of extracted features from the image. The system then automatically develops a model that correlates known data (such as spectral or spatial signatures) with targeted outputs (i.e., the features or objects of interest). The learned model then classifies and extracts the remaining targets or objects (Blundell and Opitz, 2006, p. 1)." FA *learns* inductively using an ensemble approach featuring different algorithms based on variants of artificial neural networks, decision trees, Bayesian learning, and K-nearest neighbor (Opitz and Blundell, 2008). As this software package is proprietary, details of mathematical and algorithmic background are unavailable for presentation herein. This multi-algorithm approach is thought to produce better results than the individual algorithms alone (Blundell and Opitz, 2006; Opitz and Maclin, 1999). FA outputs a classified vector layer that the analyst may then use to improve the model by selecting correctly classified features, false positives (clutter), and missed features. This information is used to re-train the model in a hierarchical fashion, meaning that classification problems are iteratively divided into increasingly smaller and more specific sub-problems (Blundell and Opitz, 2006). Initial classification errors are corrected with each subsequent iteration of this process until the analyst is satisfied with the results based on accuracy measures and visual inspection.

Neighborhood Scale Classification Methods

Xeriscaping is an increasingly popular landscaping style in the southwestern United States and other semi-arid and arid areas around the world with rapid population growth and limited water supplies. Xeric landscaping in the study area is composed mostly of gravel mulch in the front and back yards, and thus appears very similar spectrally to roads, roofs, and other impervious surfaces surrounding it. Consequently, NDVI and an edge-enhanced image were created from the four QuickBird bands as additional inputs to facilitate object-oriented classification (Figure 3). These derived products provided the object-oriented classifier with contextual information about the neighborhood of a specific pixel. However, NDVI (Equation 1) could not be used as the sole proxy for pervious areas because the spatial resolution of the QuickBird imagery was high enough to resolve single plants and small patches of turf in the otherwise gravel-covered yards. However, NDVI was still a valuable contribution to the classification process because vegetated pixels were detected by the classifier as

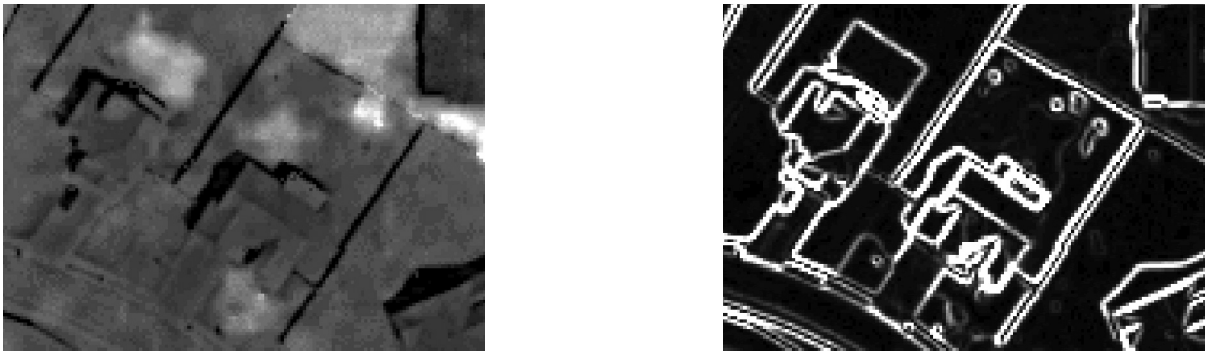

Figure 3. (a) NDVI, and (b) edge-enhanced images of typical single-home lots in the La Terraza urban watershed. These two derived input layers improved classification of impervious features relative to the panchromatic and four multispectral bands alone.

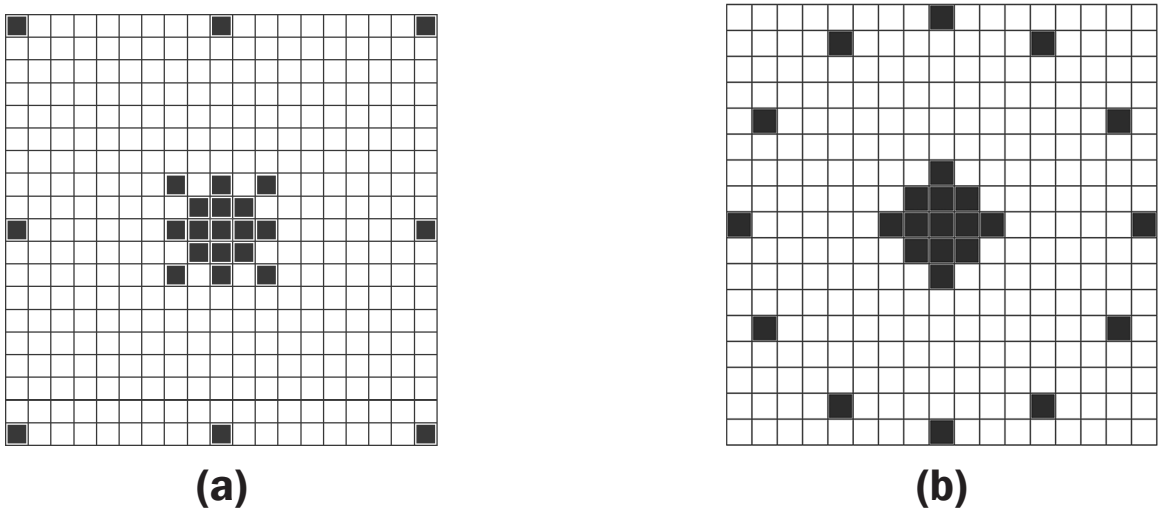

Figure 4. (a) Custom pixel pattern used by the object-oriented classification software to capture linear structures as well as smaller compact structures in the La Terraza neighborhood, and (b) "Bull's Eye 4" pixel pattern used for the regional-scale classification after it was found effective at capturing linear and square impervious features without a high amount of confusion with pervious features with similar shapes and reflectances.

pervious areas. Principal components analysis (PCA) of the four QuickBird bands was used to create four PC images. The PC bands contain the vast majority of spectral information of the sensor, thus they were included as one option to find the best possible dataset for the classification.

In order to help the classifier discriminate between impervious and pervious features in the neighborhood, an edge-enhanced image was created using the Sobel method (Kittler, 1983; Leica Geosystems, 2005), which generates horizontally and vertically edge-enhanced images before combining them into one non-directional edge-enhanced image. The Sobel filter helps to visualize the sharp linear edges of impervious features such as houses and driveways at the same time that pervious front and back yards appear with a less defined edge pattern with irregular shape. The Sobel operator is one of several edge detection algorithms such as Frei-Chen (Park, 1990 and 1999); see Davies (1984) for a survey of techniques and Sharifi *et al.* (2002) for a more recent comparison.

Three training classes, "impervious," "pervious," and "shadow" were created as vector files in an iterative process. The three vector files were combined into one training vector file.

Following preparations for the classification process, the object-oriented approach was defined by selection of which neighboring pixels are considered in the classification. Commonly, image classifiers classify a pixel based on a local window (or neighborhood), e.g., 6 × 6, of the surrounding pixels. One limitation of this approach is that it is hard to factor in the broader spatial context. In contrast, FA uses a "foveal" representation where "a learning algorithm is given a region of the image with high spatial resolution at the center (where the prediction is being made) and lower spatial resolution away from the center (Opitz and Blundell, 2008; p. 159)." By taking only the "gist" of pixels further away from the center pixel being classified, the system has less data to process and incorporates spatial context in a way more similar to human vision (Opitz and Blundell, 2008). The analyst's task is to select the particular representation (pattern of pixels) that FA will use to recognize features. There are a number of preconfigured pixel patterns in FA that work well for detecting small objects, linear objects or natural objects. However, impervious areas in La Terraza are a mix of shapes and sizes. We used FA to create a single custom pixel pattern to detect both linear features (roads and sidewalks) and smaller, compact features (houses and driveways) (Figure 4a). The inner shape of pixels was designed to discriminate compact objects, such as houses or yards. The outer points were intended to detect linear features in multiple directions. Point patterns with a higher number of points than the selected pattern, as well as patterns with fewer points and simpler patterns were less effective at discriminating features in the study area based on trial and error testing.

Three datasets were selected for the classification trials (Table 1) with the goal of determining the accuracy that could be achieved with a minimum of inputs and which combinations of inputs would yield the most accurate classification.

The *Vis* dataset resembles a basic set of high spatial resolution imagery without near infrared information, which is often available for larger areas at low cost. The other two datasets both include near infrared data as well as the additional bands in an effort to maximize the accuracy of the classifications. Although the three multispectral bands were already pan-sharpened, the panchromatic band was included in the *Pan/Vis/NIR-Plus* data set, because it covers a wide spectral range, likely enhancing patterns in the image and improving the classifier's ability to detect features.

Table 1. Overview of the Three Combinations of Input Datasets Used in the Classification Process

Dataset	Three visual bands (RGB)	Multi-spectral bands	Principal component bands	Panchromatic band	NDVI	Edge-enhanced
Vis	x					
PCA-Plus			x		x	x
Pan/Vis/NIR-Plus		x		x	x	x

The three input data set combinations were evaluated using a transformed divergence statistical separability technique to estimate which was most likely to produce the best classification. For all possible pairs of i^{th} and j^{th} training classes within each of the three groups of input data sets, a transformed divergence score (TD) is computed by comparing the degree of overlap between the probability distributions of the spectral classes of all available input bands (Richards, 2013; Singh, 1984). The score increases exponentially with increasing class distance and has a scale of 0 to 2000. TD_{ij} >1,900 represents good separability of classes *i* and *j*, scores of 1,700 to 1,900 indicate fair separability; scores below 1,700 are poor, with zero meaning classes are inseparable. The highest mean overall divergence score determines the most effective combination of inputs (Singh, 1984).

Qualitative visual analysis of classification performance was followed with statistical accuracy tests. Two hundred test points were placed using a stratified random distribution in the study area, using the manually derived pervious and impervious areas for stratification. The number of test points was determined to produce an expected classification accuracy of 85 percent and an allowable error of ±5 percent using a binomial probability approach (Jensen, 2004; Snedecor and Cochran, 1989). Based on the small GSD of the satellite imagery, photographs, and extensive knowledge of the area from fieldwork, we identified ground reference test points as pervious, impervious, or shadow. Ground reference labels and classification labels were exported into an accuracy matrix to determine total accuracy as well as errors of omission and commission and Kappa values.

Neighborhood Scale Classification Results

Based on high transformed divergence values, the *Pan/Vis/NIR-Plus* data set produced the highest separability of impervious and pervious classes (Table 2). Accordingly, the classification based on the *Pan/Vis/NIR-Plus* dataset showed the highest accuracy based on visual inspection (Figure 5c). Roads and houses were well defined and there were no substantial misclassified areas, though some clutter was still present. Misclassified areas were in the same general locations for all classifications, due likely to the variability in roof colors. The results in Table 2 show that the inclusion

Table 2. Results of the Transformed Divergence Separability Analysis, Showing How Well Input Datasets can Distinguish the Classes in the Trainings Dataset: Maximum Separability is Indicated by a Score of 2,000; Separability is Poor Below a Score of 1,700

Input Dataset	Impervious:Pervious	Impervious:Shadow	Pervious:Shadow
Vis	851	1899	1829
PCA-Plus	1618	2000	2000
Pan/Vis/NIR-Plus	1654	2000	2000

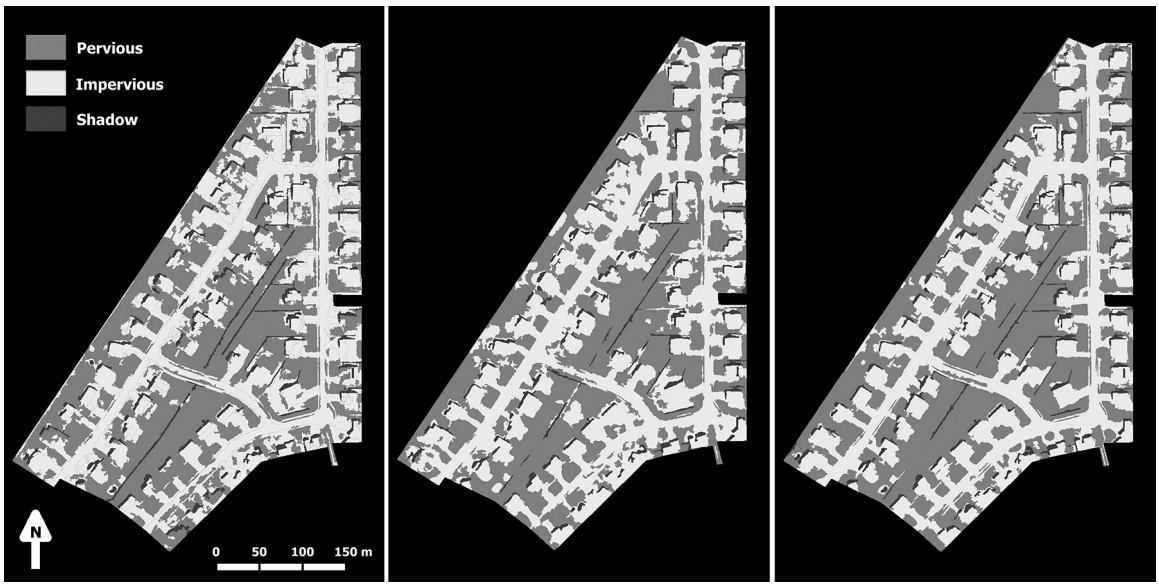

Figure 5. (a) Classification results for the *Vis* dataset, (b) *PCA-Plus* dataset, and (c) *Pan/Vis/NIR-Plus* dataset, where light gray = impervious, middle gray = pervious, and dark gray = shadow.

of the NIR band is crucial to the classification results under the specified circumstances. It also shows the usefulness of including derived layers, such as NDVI and edge-enhancements for object oriented classifications. Although a principal component analysis captures the vast majority of information of the available imagery, the addition of seemingly redundant layers (NDVI, edge-enhancement, and panchromatic) improved classification results.

Statistical tests supported both separability analyses and the visual inspection of classified imperviousness maps (Figure 6). The map created from the *Pan/Vis/NIR-Plus* dataset produced the highest overall accuracy (0.83) and Kappa value (0.68) of the three combinations of input data layers.

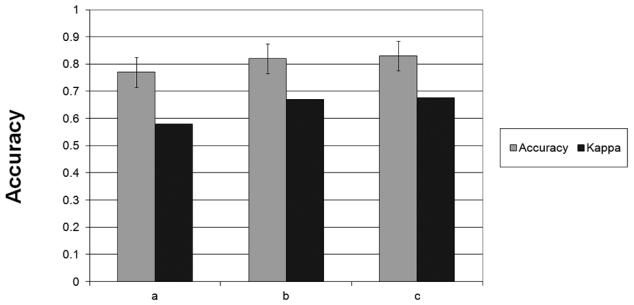

Figure 6. Comparison of total classification accuracy including confidence intervals and Kappa for all three input datasets, where a = *Vis* dataset, b = *PCA-Plus* dataset, and c = *Pan/Vis/NIR-Plus* dataset.

The manual classification result, which was used as a baseline scenario, was considered 100 percent accurate. This assumption only applies to the limited extent of the La Terraza study area, because the analyst was able to spend significant time and effort on delineating pervious and impervious areas. Applying manual delineation methods to larger areas would likely result in reduced accuracy and consistency caused by analyst fatigue, multiple analysts involved, and time and budget resource limitations.

Phase 2: Expansion of Classification Scheme to the Regional Scale

Regional Scale Classification Methods

Based on neighborhood scale classification results described in the previous section, the input data layers for the regional scale consisted of the four QuickBird multispectral bands, NDVI, edge-enhanced images, and the panchromatic band. Initial testing of the FA classifier using the seven data layers resulted in very long processing times (8 to 12 hours per scene). To make processing manageable, the images were subdivided or "diced" into smaller tiles. Dicing the QuickBird scenes resulted in a total of 21 individual image tiles to classify.

The same classification scheme of pervious, impervious, and shadow for the local scale was used to represent the regional scale as well. Two sets of training polygons were created for each image tile: impervious and shadow. Training data sets for pervious features were not created for reasons explained below.

Features considered impervious were all paved or sealed surfaces that could be discerned from inspection of the images, including parking lots, sidewalks, airport runways, buildings, residences, driveways, and recreational areas with sealed surfaces such as running tracks. A few basic types of roads could be distinguished using contextual clues and ancillary data such as Google™ StreetView street-level photographs: dark asphalt roads, non-asphalt paved roads, graded but unpaved roads, and dirt roads. The classifier was trained to recognize the first two types as impervious. Everything that was not paved or sealed was considered pervious. Pervious land-covers consisted mostly of desert scrub vegetation, undeveloped bare soil, unpaved roads and trails, agricultural plots, trees, playing fields, lawns, parks, and road medians.

The training polygons for shadows were designed primarily to capture the shadows cast by buildings and structures and shadows associated with natural surfaces, e.g., from tall trees in forested areas. In the southwestern portion of the study area with a relatively high amount of relief, large areas were cast in shadow by ridgelines. However, these areas of rugged topography were clearly undeveloped, so we generally limited the final training data sets to those shadows cast by individual features only, and not entire areas of pervious cover.

Extensive testing and examination of the results of different combinations of classifier settings and training data sets helped determine an effective classification model that could be applied across all tiles. This process retained most of the settings used for the neighborhood scale but also deviated in some ways due to the far greater spatial extent and diversity of features that had to be accounted for in order to optimize the accuracy of results.

Initial classification trials using the same settings as the neighborhood scale analysis resulted in confusion of certain bare natural surfaces with impervious surfaces that had very similar shapes and reflectances. The result was the erroneous detection or "false positives" of impervious shapes in natural areas with bare patches, which littered the undeveloped areas of the images. This effect was minimized by altering the classification method such that rather than using a "wall-to-wall" classification that classifies all pixels into three input training data classes (shadow, impervious, and pervious) as for La Terraza, only two were used (shadow and impervious), leaving the remainder of the image temporarily unclassified background pixels. Testing results indicated that the use of two training layers minimized the confusion between pervious and impervious training polygons, which seemed to be the primary source of this type of error. The reasoning behind the altered classification scheme was that pixels that were neither shadow nor impervious (the "background" pixels) were pervious, therefore the two input training classes were used to extract just those two classes initially, and the unclassified background areas (the pervious areas) could then be classified in a later step and renamed "pervious."

The input representation settings determine the pattern used by the classifier to detect impervious features. We found based on trial and error testing that the "Bull's Eye 4" pixel pattern with a width of 17 pixels (Figure 4b) was more consistently effective than others at detecting a wide range of both linear and rectangular/square impervious features with relatively low amounts of misclassified false positive clutter in pervious areas. The minimum object size was set to 100 pixels so that objects smaller than 100 pixels would be aggregated with neighboring features after classification. As for Miller *et al.* (2009), this setting seemed to produce the most optimal balance of omission and commission errors.

The classifier model was applied to 18 of the 21 tiles using two-class (impervious and shadow) training polygon sets. Three tiles (321, 411, and 6) contained so few impervious surfaces that it made more sense to classify them manually. The output from the classification process for each image tile was a two-class vector layer of impervious and shadow.

One of the main advantages of the FA classifier is the array of tools used to iteratively refine the classifier until the desired result is achieved. This process involved splitting the classified output vector layer into separate layers to be cleaned up individually as necessary based on visual inspection of the results. One set of tools works by selecting polygons and digitizing portions of features that were correctly

and incorrectly classified in order to re-train the classifier more precisely. Another set of tools requires the user to draw polygons to capture examples of missing features. These tools were used to iteratively retrain the classifier until a satisfactory classification was produced, i.e., based on visual confirmation of a relatively minimal amount of clutter, no obvious systematic or major misclassifications, and generally good capturing of impervious objects throughout the image.

Finally, the cleaned-up impervious and shadow layers were recombined, with settings adjusted such that any pixels not included in either of the two classes during the union of layers are classified into polygons. These now-classified "background" polygons, i.e., the remainder of the image that was initially left unclassified, formed the pervious class. The resulting three-class polygon layer was converted to raster format and pixel values were reclassified as necessary into 1 (impervious), 2 (pervious), and 3 (shadow) for consistency across tiles.

The accuracy of each classified image tile was assessed using a stratified random sampling approach with a minimum per-class sample size of 50 points. The impervious class was intentionally over-sampled because discriminating impervious surfaces was the primary goal of the classification. Since the small impervious areas in the three manually classified tiles were easily digitized and all remaining area was impervious, it was not necessary to create a shadow class.

In-field validation of sample points was not feasible due to the large number of points needed to evaluate the accuracy of all 21 tiles. Instead a combination of aerial photos, field photographs of locations on the ground, and experiential knowledge from visits to the La Terraza subdivision was used. These ancillary data helped distinguish similar surfaces such as pervious gravel and xeriscaped yards from fully impervious roads, parking lots, and roofs. There were 5,257 sample points total for all 21 scenes combined, a much larger sample than a recent study using a similar amount of QuickBird data and 1,735 validation points (Campos *et al.*, 2010).

Error matrices were computed for each classified image tile (example shown in Table 3) and used to calculate overall accuracy and producer's and user's accuracies of each individual class (Table 4). The Kappa coefficient of agreement was also calculated for each overall classification and for individual classes (Table 5).

Table 3. Sample Error Matrix for Classified Image Tile 221. Error Matrices were Used to Derive Accuracy Assessments for Each Tile

	Predicted Class			
Actual Class	**Impervious**	**Pervious**	**Shadow**	**Row Total**
Impervious	**75**	25	0	100
Pervious	2	**96**	2	100
Shadow	10	0	**65**	75
Column Total	87	121	67	**236**

In order to produce a final classified image compatible with hydrologic modeling applications, all shadow pixels were reclassified as either pervious or impervious. A neighborhood analysis tool was used to reassign the value of all shadow pixels based on the majority value of the surrounding pixels in the 3 × 3 "neighborhood." This process was run on every classified image that had any shadow pixels, i.e. not classified manually, producing a reclassified set of image tiles with only two classes, impervious and pervious. Finally, the 21 classified image tiles were mosaicked together.

Regional Scale Classification Results

The final regional mosaicked image had a total area of 1,179 km^2. Based on this image, the study area contains an estimated 18.6 km^2 of impervious surfaces, or 1.6 percent of the total area, mainly concentrated in the highly developed core of Sierra Vista, and to a lesser extent Ft. Huachuca.

Table 4. Per-class Classification Accuracies Grouped by Degree of Imperviousness as a Percentage of Total Area of the Image Tile: N/A indicates tile was classified manually and no shadow pixels were classified

Imperviousness Grouping	Image Tile ID	Overall Accuracy (%)	Impervious Producer's Accuracy (%)	Impervious User's Accuracy (%)	Pervious Producer's Accuracy (%)	Pervious User's Accuracy (%)	Shadow Producer's Accuracy (%)	Shadow User's Accuracy (%)
Most Developed (3-7% Impervious)	211	91.2	88.4	95.0	92.2	94.0	98.2	73.3
	221	85.8	86.2	75.0	79.3	96.0	97.0	86.7
	512	89.1	94.6	87.0	84.6	93.0	89.0	86.7
	212	88.4	91.0	81.0	87.4	90.0	86.8	96.0
More Developed (1-2% Impervious)	422	91.3	90.0	90.0	87.6	99.0	100.0	82.7
	222	83.6	91.4	74.0	72.4	97.0	98.3	78.7
	312	89.8	87.1	88.0	86.5	96.0	100.0	84.0
	121	90.9	91.6	87.0	85.8	97.0	98.5	88.0
	412	87.3	96.2	75.0	76.3	100.0	98.5	86.7
	311	82.9	91.0	71.0	73.3	96.0	92.4	81.3
	322	86.6	87.1	81.0	80.5	99.0	98.3	77.3
Least Developed (<1% Impervious)	111	89.8	98.8	82.0	79.4	100.0	98.5	86.7
	112	84.4	91.5	72.0	77.5	100.0	97.3	72.0
	421	85.8	100.0	74.7	76.3	100.0	97.4	74.0
	511	89.0	100.0	70.0	84.8	95.0	90.6	96.0
	411	99.4	100.0	98.7	99.0	100.0	n/a	n/a
	122	96.0	100.0	96.0	92.6	100.0	100.0	88.0
	522	89.0	93.6	88.0	87.9	94.0	87.0	80.0
	521	85.0	100.0	74.0	84.3	86.0	77.1	94.0
	321	100.0	100.0	100.0	100.0	100.0	n/a	n/a
	6	99.4	100.0	98.7	99.0	100.0	n/a	n/a

Table 5. Overall and Per-class Kappa Coefficients (κ) with Confidence Intervals for Overall κ: N/A indicates tile was classified manually and no shadow pixels were classified

Imperviousness Grouping	Image Tile ID	Overall k	Impervious k	Pervious k	Shadow k	95 % C.I. upper	95 % C.I. lower
Most Developed (3-7 % Impervious)	211	0.86	0.91	0.89	0.70	0.90	0.81
	221	0.78	0.63	0.93	0.82	0.83	0.74
	512	0.83	0.80	0.88	0.82	0.89	0.78
	212	0.82	0.72	0.84	0.94	0.88	0.77
More Developed (1-2 % Impervious)	422	0.87	0.84	0.98	0.78	0.92	0.82
	222	0.75	0.63	0.94	0.73	0.82	0.69
	312	0.85	0.81	0.93	0.79	0.90	0.79
	121	0.86	0.80	0.95	0.84	0.91	0.81
	412	0.81	0.65	1.00	0.82	0.86	0.75
	311	0.74	0.60	0.92	0.75	0.81	0.67
	322	0.79	0.71	0.98	0.71	0.85	0.74
Least Developed (< 1 % Impervious)	111	0.85	0.74	1.00	0.82	0.90	0.79
	112	0.75	0.62	1.00	0.66	0.81	0.69
	421	0.77	0.66	1.00	0.69	0.84	0.70
	511	0.82	0.64	0.89	0.95	0.89	0.75
	411	0.99	0.98	1.00	n/a	1.07	0.91
	122	0.94	0.95	1.00	0.85	1.00	0.87
	522	0.82	0.84	0.87	0.74	0.89	0.75
	521	0.76	0.68	0.71	0.91	0.84	0.68
	321	1.00	1.00	1.00	n/a	1.08	0.92
	6	0.99	0.98	1.00	n/a	1.07	0.91

Mean overall classification accuracy for the 21 individual tiles was 89.7 percent with a range of 82.9 to 100 percent (Table 4). 100 percent accuracy was associated only with those images that were classified manually. Excluding those three images, the mean overall accuracy was 88.1 percent, which was not significantly lower (t = 1.213; p >0.05; df = 34). Mean producer's accuracy for the impervious class (94 percent) was higher than for the pervious class (85 percent) (t = 4.344; p <0.05; df = 33). Mean user's accuracy for the impervious class was accurate at 84 percent, but significantly lower than the pervious class (97 percent) (t = -5.568; p <0.05; df = 26). The mean Kappa coefficient was higher for the pervious class than the impervious class (t = -5.087; p <0.05; df = 31).

Limitations of the Regional Scale Classification

Limitations of computing power necessitated subdividing the original images into separate tiles. Although the classification accuracies were satisfactory, the fact that each image tile was classified based on different training data unavoidably resulted in some discrepancies in accuracy among the tiles, which was evident from visual inspection. For example, roads in one image tile may be more conservatively classified than in another, resulting in edge effects in the final mosaic where the road is wider on one side of a cutline than on the other. Had it been possible to classify the entire study area at once, only one training data set would have been used and edge effects would not have resulted. Inconsistencies among classifications such as edge effects thus stem in part from limitations of processing power.

In each classified image tile, the goal was to strike a balance between accuracy and efficiency in terms of user time spent (e.g., preparing and iteratively revising training polygons) and computer processing time. With the iterative revising tools within FA it is possible to eventually achieve a high level of accuracy. This often required a significant time investment by the user, though it should be noted that during the regional scale phase a newer version of FA was released with the capability of running test classifications on any portion of a scene in a matter of minutes. This likely did not have a meaningful impact on classification accuracy compared to the previous version, but did significantly speed up testing of the effectiveness of training polygons. Further, after a few iterations of re-training the classifier, a point of diminishing returns was reached where the gains in additional accuracy (based on visual inspection and comparison) were relatively small compared to the amount of time required to achieve them. Each image tile had its own unique set of features, and consequently this point of diminishing returns was different for each tile. Since a perfect classification is impossible, in each tile some tradeoffs had to be made. For instance, in several cases achieving an extremely high degree of accuracy for roads in a given tile was not possible without producing a large number of false positives for those natural surfaces with similar spectral and shape characteristics to impervious roads, which would lead to lower user's accuracy (greater errors of commission). In this case, creating a classification model that produced a balance between the two, i.e., adequate classification of roads without a high amount of false positives in pervious areas, meant necessarily sacrificing some accuracy for roads. Discrepancies that were apparent from visual inspection in the final image are the result of these unique tradeoffs that had to be made in each tile based on the variety of features that it contained. In sum, each classified tile reflects the best effort of the user to achieve a satisfactory level of accuracy based on error matrices within a reasonable amount of time invested in refining the initial classification.

Sources of Confusion between Classes and Discrepancies in Accuracy in the Regional Scale Classification

After examining the initial classification result, an important source of confusion was found between rectangular building structures with high reflectances and pervious (although likely highly compacted) natural surfaces with similar reflectances and shapes. In many cases, it was evident from ancillary data that the natural surfaces were impervious rock. In this case, during accuracy assessment we counted it as correct if it was classified as impervious. There were also instances, however, where the surface was not rock, but rather brightly-reflecting bare soil or a graded unpaved road with similar rectangular shape to buildings. When these areas were classified as impervious, they were considered a misclassification. Cleanup tools in FA were useful for removing this misclassification to a significant degree. Additionally, roads and other surfaces are not spectrally homogeneous, and a training polygon in one part of a road may not be adequate to make the classifier capture all parts of a road feature. Training data sets took this into consideration, but could not always prevent gaps in road features. In many cases, creating a training polygon around one portion of a road led to more false positives in other spectrally similar non-road objects. When this limitation was found, we generally tried to find an optimal balance between errors of omission and commission. Some amount of this type of error was virtually unavoidable, but could be improved using the hierarchical iterative learning tools described previously.

Kappa values (Table 5) indicate that the classifier was more accurate for the pervious classes than the impervious classes, although impervious Kappa values were mostly very good, with a mean value of 0.77. In all tiles, the Kappa values show that overall and per-class accuracies were not due to chance. Percentage accuracy results show that while producer's accuracy was quite high for impervious classes, the main source of error was found in the user's accuracy. In other words, sample pixels that were actually impervious were classified at a high level of accuracy, while those pixels that were not impervious were more frequently classified incorrectly as impervious. Visual inspection of the classified results showed that this is to some degree related to the inability of the classifier to distinguish boundaries and edges between certain adjacent features. For example, in many cases the roads outside of the highly developed core of Sierra Vista appeared in the imagery to have a somewhat fuzzy gradient from the paved surface to the shoulder to the undeveloped ground. This is related to the lack of vegetation along the sides of roads and in medians, which is typical for this region but less common in humid regions where grass and trees often create distinct boundaries between paved surfaces. Another source of error stemmed from confusion between bright road stripes and unpaved, non-vegetated medians. These two types of features are sometimes very similar in shape and spectral signature in this region, posing a challenge to creating training datasets that can produce a separation of these features in the classification.

Finally, this study required the classification of all impervious features, regardless of shape. This required a pixel pattern that was effective for detecting a wide variety of features. Had the task been to classify only a subset of impervious features, e.g., roads, a more specialized pixel pattern could have been used, likely improving accuracy.

Conclusions and Recommendations

Typical methods for detecting and mapping impervious surfaces in arid and semiarid environments using only vegetation indices are not ideal due to sparseness of green vegetation cover and the presence of xeriscaping and senescent, non-green vegetation, making detecting impervious surfaces in arid and semi-arid environments challenging. Object-oriented classification algorithms in combination with carefully selected input data offer a solution by incorporating not only spectral data, but also information about the pixel environment, such as patterns and neighborhood relations. This study demonstrated at the scale of an individual neighborhood that the automated, object-oriented classification of QuickBird data to create imperviousness maps of semi-arid urban areas produced results with reasonable accuracy that can be obtained much more efficiently and in an objectively repeatable manner than using traditional manual delineation techniques. The approach used for the neighborhood classification was then adapted to produce a map of impervious surfaces in the entire city of Sierra Vista, Arizona and the surrounding sub-watershed.

This method demonstrated a common limitation of high spatial resolution imagery, which is that broad spatial coverage can be difficult due to the high amount of data per image. Using multiple QuickBird images created a very large amount of data to process. If the images must be subdivided and classified individually for this reason, achieving a high level of consistency from tile to tile may present a challenge for the user and likely increases the chances of user error associated with performing many classifications.

Three recommendations are made for using the impervious surface classification method reported here.

1. This study presents the difficulties involved in mapping impervious areas in semi-arid and arid urban environments using currently available sensor and classification technology. A key finding is the importance of NIR information to conduct this classification. Aerial imagery without NIR bands is frequently available for urban areas; however, this type of imagery would not be suited to classify impervious areas using the proposed approach.
2. The proposed method should be applied to areas with uniform land-cover. Conducting the classification for the whole Sierra Vista watershed, including urban, rural, and mountainous areas, resulted in decreased efficiency of the training algorithm. Based on this study, we recommend applying the proposed method to urban areas, where imagery with small GSD is most beneficial to capture the rapid succession of pervious and impervious areas. Imagery from sensors with larger GSD, which typically have more spectral information appear more suitable for less populated areas.
3. While object-oriented classification using high-resolution data was labor- and cost-effective when applied at the scale of a relatively small region (less than 1,200 km^2 in this case), it would be increasingly less so the larger the region. Coarser resolution imagery such as Landsat would still be more appropriate for assessing imperviousness for bigger regions such as the Baltimore, Maryland - Washington, D.C. metro corridor (Sexton *et al.*, 2013). Although it was not a goal of this study, it would be a worthwhile future task to quantify more precisely the size area at which the use of high-resolution imagery for mapping impervious surfaces becomes prohibitive and more appropriate for coarser-resolution imagery.

Despite its limitations, this classification approach should enable users to match acquisition and analysis of satellite data to census numbers and thus repeat the method in synchrony with future censuses or soil and land surveys in order to track increases in impervious surfaces associated with the growth of the area over time. Relating impervious surface changes to longitudinal hydrologic data will lead to a more accurate holistic understanding of the potential impacts of continued

growth on storm water runoff and utilization of this water resource for groundwater recharge. We suggest that the methods described here be further developed through application in other rapidly growing areas in the Western US and other arid and semiarid environments.

Acknowledgments

The Upper San Pedro Partnership and the USDA-Agricultural Research Service supported this research, and they are gratefully acknowledged. Thanks are extended to Dr. Stuart Marsh for methodological advice. Chandra Holifield-Collins is also acknowledged for coordinating the acquisition of the imagery used in this analysis.

References

Arnold, Jr., C.L., and C.J. Gibbons, 1996. Impervious surface coverage: The emergence of a key environmental indicator, *Journal of the American Planning Association*, 62(2):243–258.

Bauer, M.E., J.K. Doyle, and N.J. Heinert, 2002. Impervious surface mapping using satellite remote sensing, *Proceedings of the Geoscience and Remote Sensing Symposium, 2002. IGARSS'02, 2002 IEEE International*, pp. 2334–2336.

Blaschke, T., S. Lang, and G.J. Hay, 2008. *Object-based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications*, Springer, Berlin, 817 p.

Blundell, J., and D. Opitz, 2006. Object recognition and feature extraction from imagery: The Feature Analyst® approach, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36:4.

Campos, N., R. Lawrence, B. McGlynn, and K. Gardner, 2010. Effects of LiDAR-Quickbird fusion on object-oriented classification of mountain resort development, *Journal of Applied Remote Sensing*, 4(043556):043556.

Chormanski, J., T. Van de Voorde, T. De Roeck, O. Batelaan, and F. Canters, 2008. Improving distributed runoff prediction in urbanized catchments with remote sensing based estimates of impervious surface cover, *Sensors*, 8(2):910–932.

Colby, B.G., and K. L. Jacobs, 2007. *Arizona Water Policy: Management Innovations in an Urbanizing, Arid Region*, RFF Press, Washington, D.C., 247 p.

Davies, E.R. 1984. Circularity - A new principle underlying the design of accurate edge orientation operators, *Image and Vision Computing*, 2(3):134–142.

Goodrich, D.C., A. Chehbouni, B. Goff, B. MacNish, T. Maddock, S. Moran, W.J. Shuttleworth, D.G. Williams, C. Watts, and L.H. Hipps, 2000. Preface paper to the Semi-Arid Land-Surface-Atmosphere (SALSA) program special issue, *Agricultural and Forest Meteorology*, 105(1):3–20.

Huete, A., 1988. A soil-adjusted vegetation index (SAVI), *Remote Sensing of Environment*, 25(3):295–309.

Jensen, J.R., 2004. *Introductory Digital Image Processing,* Third edition., Pearson Prentice Hall, Upper Saddle River, New Jersey, 544 p.

Kennedy, J.R., D.C. Goodrich, and C.L. Unkrich, 2013. Using the KINEROS2 modeling framework to evaluate the increase in storm runoff from residential development in a semi-arid environment, *Journal of Hydrologic Engineering*, 18(6):698-706

Kittler, J., 1983. On the accuracy of the Sobel edge detector, *Image and Vision Computing*, 1(1):37–42.

Kressler, F.P., T.B. Bauer, and K.T. Steinnocher, 2001. Object-oriented per-parcel land use classification of very high resolution images, *Remote Sensing and Data Fusion over Urban Areas*, *Proceedings of the IEEE/ISPRS Joint Workshop 2001*, pp. 164–167.

Lang, S., 2008. Object-based image analysis for remote sensing applications: Modeling reality – Dealing with complexity, *Object-based Image Analysis: Spatial Concepts for Knowledge-driven Remote Sensing Applications* (T. Blaschke, S. Lang, and G.J. Hay, editors), Springer, Berlin, pp. 3–27.

Lee, J.G., and J.P. Heaney, 2003. Estimation of urban imperviousness and its impacts on storm water systems, *Journal of Water Resources Planning and Management*, 129(5):419–426.

Leica Geosystems, 2005. *ERDAS Field Guide*.

McMahon, G., 2007. Consequences of land-cover misclassification in models of impervious surface, *Photogrammetric Engineering & Remote Sensing*, 73(12):1343–1353.

Miller, J.E., G.R. Hess, and C.E. Moorman, 2007. Southern two-lined salamanders in urbanizing watersheds, *Urban Ecosystems*, 10(1):73–85.

Miller, J.E., S.A.C. Nelson, and G.R. Hess, 2009. An object extraction approach for impervious surface classification with very-high-resolution imagery, *The Professional Geographer*, 61(2):250–264.

Opitz, D., and S. Blundell, 2008. Object recognition and image segmentation: The Feature Analyst® approach, *Object-based Image Analysis: Spatial Concepts for Knowledge-driven Remote Sensing Applications* (T. Blaschke, S. Lang, and G.J. Hay, editors), Springer, Berlin, pp. 153–167.

Opitz, D., and R. Maclin, 1999. Popular ensemble methods: An empirical study, *Journal of Artificial Intelligence Research*, 11:169–198.

Park, R.H., 1990. A fourier interpretation of the Frei-Chen edge masks, *Pattern Recognition Letters*, 11(9):631–636.

Park, R.H., 1999. Interpretation of eight-point discrete cosine and sine transforms as 3 × 3 orthogonal edge masks in terms of the Frei–Chen masks, *Pattern Recognition Letters*, 20(8):807–811.

Platt, R.V., and L. Rapoza, 2008. An evaluation of an object-oriented paradigm for land use/land cover classification, *The Professional Geographer*, 60(1):87–100.

Richards, J.A., 2013. *Remote Sensing Digital Image Analysis: An Introduction*, Springer, Berlin, 494 p.

Sawaya, K.E., L.G. Olmanson, N.J. Heinert, P.L. Brezonik, and M.E. Bauer, 2003. Extending satellite remote sensing to local scales: Land and water resource monitoring using high-resolution imagery, *Remote Sensing of Environment*, 88(1):144–156.

Sexton, J.O., X. Song, C. Huang, S. Channan, M.E. Baker, and J.R. Townshend, 2013. Urban growth of the Washington, D.C. - Baltimore, MD metropolitan region from 1984 to 2010 by annual, Landsat-based estimates of impervious cover, *Remote Sensing of Environment*, 129:42–53.

Sharifi, M., M. Fathy, and M.T. Mahmoudi, 2002. A classified and comparative study of edge detection algorithms, *Proceedings of the International Conference on Information Technology: Coding and Computing, 2002*, pp. 117–120.

Shuster, W.D., J. Bonta, H. Thurston, E. Warnemuende, and D.R. Smith, 2005. Impacts of impervious surface on watershed hydrology: A review, *Urban Water Journal*, 2(4):263–275.

Singh, A., 1984. Some clarifications about the pairwise divergence measure in remote sensing, *International Journal of Remote Sensing*, 5(3):623–627.

Slonecker, E.T., D.B. Jennings, and D. Garofalo, 2001. Remote sensing of impervious surfaces: A review, *Remote Sensing Reviews*, 20(3):227–255.

Snedecor, G.W., and W.G. Cochran, 1989. *Statistical Methods*, Eighth edition, Iowa State University Press, Ames, Iowa, 503 p.

Thanapura, P., D.L. Helder, S. Burckhard, E. Warmath, M. O'Neill, D. Galster, 2007. Mapping urban land cover using QuickBird NDVI and GIS spatial modeling for runoff coefficient determination, *Photogrammetric Engineering & Remote Sensing*, 73(1):57–65.

Thomas, N., C. Hendrix, and R.G. Congalton, 2003. A comparison of urban mapping methods using high-resolution digital imagery, *Photogrammetric Engineering & Remote Sensing*, 69(9):963–972.

Tsai, Y.H., D. Stow, and J. Weeks, 2011. Comparison of object-based image analysis approaches to mapping new buildings in Accra, Ghana using multi-temporal QuickBird satellite imagery, *Remote Sensing*, 3(12):2707–2726.

Vias, A.C., and J.I. Carruthers, 2005. Regional development and land use change in the Rocky Mountain west, 1982-1997, *Growth and Change*, 36(2):244–272.

Weng, Q., 2007. *Remote Sensing of Impervious Surfaces*, CRC Press, Boca Raton, Florida, 488 p.

Weng, Q., 2012. Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends, *Remote Sensing of Environment*, 117:34–49.