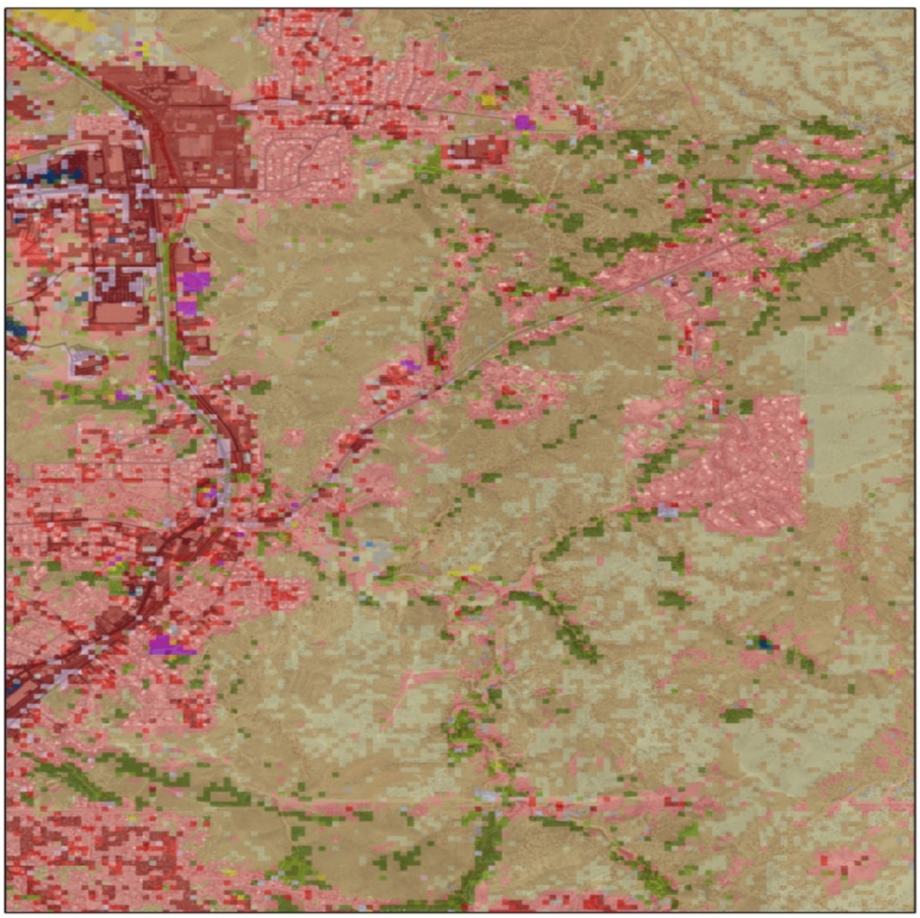
Prepared in Cooperation with The University of Arizona, School of Natural Resources and the Environment

A Multitemporal (1979-2009) Land-Use/Land-Cover Dataset of the Binational Santa Cruz Watershed

By Miguel L. Villarreal, Laura M. Norman, Cynthia S.A. Wallace, and Charles van Riper III



Open-File Report 2011–1131

U.S. Department of the Interior
U.S. Geological Survey

U.S. Department of the Interior

KEN SALAZAR, Secretary

U.S. Geological Survey

Marcia K. McNutt, Director

U.S. Geological Survey, Reston, Virginia: 2011

For product and ordering information:
World Wide Web: *http://www.usgs.gov/pubprod*
Telephone: 1-888-ASK-USGS

For more information on the USGS—the Federal source for science about the Earth,
its natural and living resources, natural hazards, and the environment:
World Wide Web: *http://www.usgs.gov*
Telephone: 1-888-ASK-USGS

Suggested citation:

Villarreal, M.L., Norman, L.M., Wallace, C.S.A., and van Riper, Charles, III, 2011, A multitemporal (1979–2009) land-use/land-cover dataset of the binational Santa Cruz watershed: U.S. Geological Survey Open-File Report 2011-1131, 26 p. [http://pubs.usgs.gov/of/2011/1131/].

Contents

Abstract 1
Introduction 2
Study Area 2
Methods 4
Image Acquisition 4
Atmospheric and Radiometric Calibration 6
Image Transformations and Ancillary Datasets 6
Training Data 7
Land-Cover Class Descriptions (From NLCD 2001) 8
Classification and Regression Tree Model 10
Post-Processing 10
Accuracy Assessment 11
Results 12
Accuracy Results 17
Conclusions 19
Acknowledgments 24
References Cited 24
Appendix. Multitemporal Tasseled Cap Coefficients for Landsat MSS Data 26

Figures

1. Map of the upper Santa Cruz Watershed 4
2. Random sample of USGS quadrangles used to train and assess the accuracy of LULC in the SCW 8

3. Example of data issues with original Landsat scene (left) and related classification errors (right). These errors were corrected manually by recoding the raster with the correct land cover types as determined from aerial photography 11

4. 1979 Land Use/Land Cover map 13

5. 1989 Land Use/Land Cover map 14

6. 1999 Land Use/Land Cover map 15

7. 2009 Land Use/Land Cover map 16

Tables

1. Landsat MSS and Landsat TM images used to develop Land Use/ Land Cover maps 5

2. Class area 1979-2009 (hectares 17

3. Error matrix for 1979 LULC map 20

4. Error matrix for 1989 LULC map 21

5. Error matrix for 1999 LULC map 22

6. Error matrix for 2009 LULC map 23

This page intentionally left blank.

A Multitemporal (1979–2009) Land-Use/Land-Cover Dataset of the Binational Santa Cruz Watershed

By Miguel L. Villarreal[1], Laura M. Norman[2], Cynthia S.A. Wallace[2], and Charles van Riper III[3]

Abstract

Trends derived from multitemporal land-cover data can be used to make informed land management decisions and to help managers model future change scenarios. We developed a multitemporal land-use/land-cover dataset for the binational Santa Cruz watershed of southern Arizona, United States, and northern Sonora, Mexico by creating a series of land-cover maps at decadal intervals (1979, 1989, 1999, and 2009) using Landsat Multispectral Scanner and Thematic Mapper data and a classification and regression tree classifier. The classification model exploited phenological changes of different land-cover spectral signatures through the use of biseasonal imagery collected during the (dry) early summer and (wet) late summer following rains from the North American monsoon. Landsat images were corrected to remove atmospheric influences, and the data were converted from raw digital numbers to surface reflectance values. The 14-class land-cover classification scheme is based on the 2001 National Land Cover Database with a focus on "Developed" land-use classes and riverine "Forest" and "Wetlands" cover classes required for specific watershed models. The classification procedure included the creation of several image-derived and topographic variables, including digital elevation model derivatives, image variance, and multitemporal Kauth-Thomas transformations. The accuracy of the land-cover maps was assessed using a random-stratified sampling design, reference aerial photography, and digital imagery. This showed high accuracy results, with kappa values (the statistical measure of agreement between map and reference data) ranging from 0.80 to 0.85.

---

[1] University of Arizona, School of Natural Resources and the Environment, Tucson, AZ 85721.

[2] U.S. Geological Survey, 520 N. Park Avenue, Tucson, AZ 85719-5035.

[3] U.S. Geological Survey, Southwest Biological Science Center SDRS, School of Natural Resources and the Environment, 1110 E. South Campus Dr., Room 123, University of Arizona Tucson, AZ 85721-0033.

Introduction

The National Land Cover Datasets of 1992 (NLCD) and 2001 (NLCD 2001) consist of 30-m resolution thematic data representing the land use/land cover (LULC) of the conterminous United States (Vogelmann and others, 2001; Homer and others, 2007). The datasets have proven invaluable for many wide-ranging scientific research projects in the United States (Riitters and Wickham, 2003; Scanlon and others, 2005; Schulte and others, 2007). One limitation of the NLCD is the spatial extent, which is limited to the conterminous United States, making it difficult to evaluate LULC change dynamics in watersheds that cross national political boundaries. We developed an NLCD-style LULC classification for the upper Santa Cruz watershed of southern Arizona, United States, and northern Sonora, Mexico, to support new and ongoing research requiring detailed, multitemporal LULC data (Norman and others, 2010; Norman, L.M., and others, unpublished draft report on mapping human well-being in the U.S.-Mexico borderlands.).

The LULC classification process that we employed involved the following steps:

1. Acquire Landsat imagery and ancillary data sets of the study area.
2. Correct and calibrate raw satellite imagery to remove atmospheric influences.
3. Create spectral transformations from corrected imagery and spatial derivatives from elevation and hydrologic data sets.
4. Develop multiple training data sets representing 14 land use and land cover classes over the 30-year period.
5. Create classified LULC maps using Classification and Regression Tree (CART) models, spatial data sets, and training data.
6. Assess the accuracy of LULC maps using high-resolution aerial imagery.

The following sections of this report provide details of our classification procedure and accuracy assessment of Santa Cruz watershed LULC maps.

Study Area

The 9,146-km^2 upper Santa Cruz watershed (SCW) in southern Arizona, United States, and northern Sonora, Mexico (fig. 1), contains varying topography, vegetation, and land-use patterns. The Santa Cruz River begins in the San Rafael Valley of Arizona, flows south into Mexico, crosses again into the United States near “Ambos Nogales” (Nogales, Sonora, and Nogales, Arizona), eventually joining the Gila River near Phoenix, Arizona. The upper Santa Cruz watershed contains two major urbanized areas, the Tucson metropolitan area and Ambos Nogales, surrounded by various vegetation communities ranging from desert grasslands to mixed-conifer forests. Elevation of the watershed ranges from 628 to 2,798 m. The SCW is characterized by mild winter and high summer temperatures and a bimodal precipitation pattern, with more than half of the precipitation contributed in mid-summer by the North American monsoon and the remainder falling in winter from Pacific frontal storms. Average annual precipitation recorded at Tumacácori National Historical Park (NHP) between 1946 and 2005 was 38 cm (Western Regional Climate Center, http://www.wrcc.dri.edu/*)*. We delineated the study area watershed boundary by using a combination of known 8-digit Hydrologic Unit Code (HUC) boundaries merged with Instituto Nacional de Estadística Geografíca e Informática (INEGI) watershed boundaries and a 30-m digital elevation model (DEM) (Norman and others, 2010).

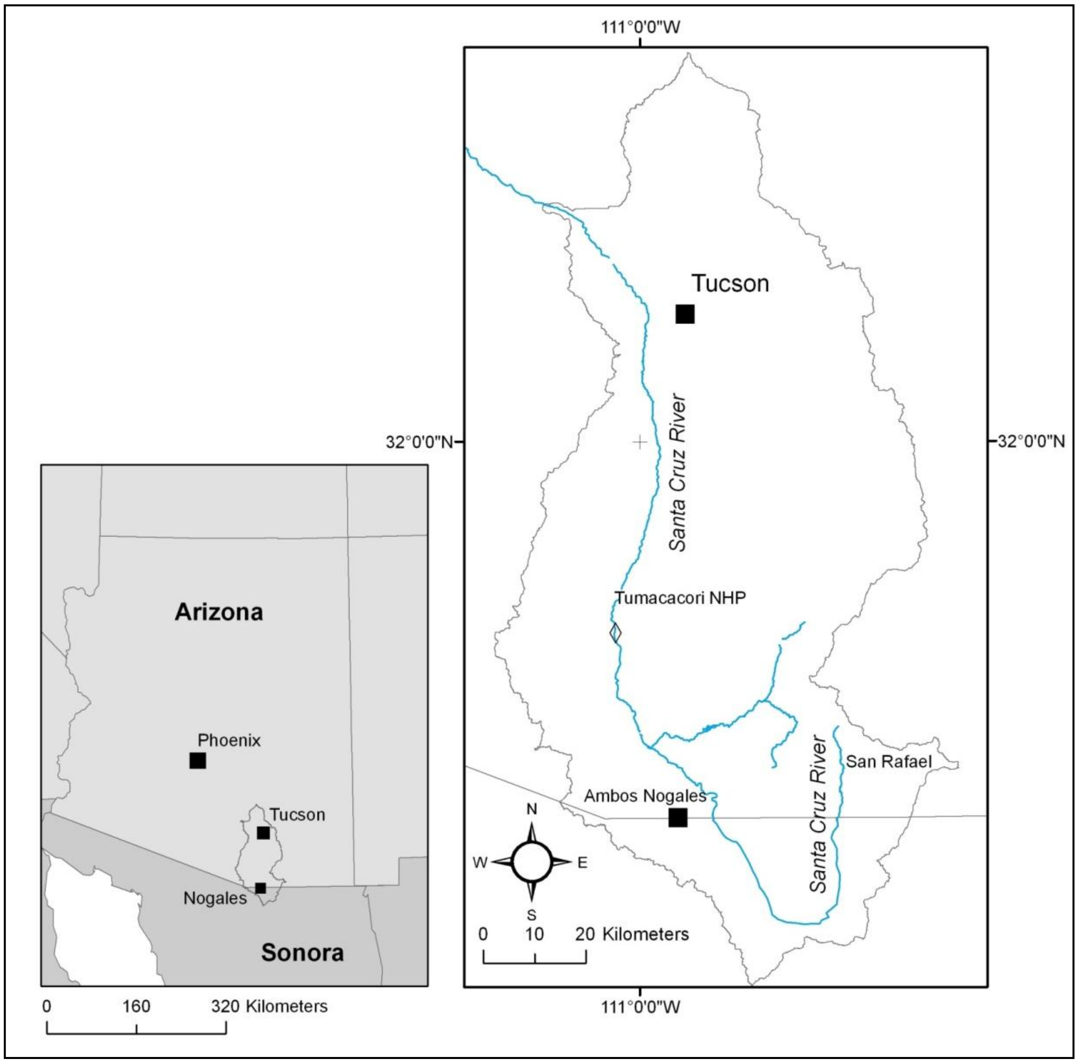


Figure 1. Map showing the upper Santa Cruz watershed boundary, urban areas, Tumacácori National Historical Park, and the Santa Cruz River and its headwaters.

Methods

Image Acquisition

Twenty-two orthorectified satellite images were acquired from glovis (http://glovis.usgs.gov) (table 1). Because the SCW is large, a single land-cover map required multiple scenes from multiple path/rows (P/R): Landsat MSS P38/R37 and P38/R38, and Landsat TM P35/R38, P36/R37, and P36/R38. Of the 22 scenes, 4 were Landsat MSS scenes from 1979, and 18 were Landsat TM scenes from 1989, 1999, and 2009. Biseasonal Landsat images were collected in June and August/September of each year; one from the dry early summer and one following the “green-up” that occurs two months later with rain from the

North American monsoon. Biseasonal images for each classification date were selected as close to anniversary as possible (table 1). All Landsat TM bands were used in the classification except for the thermal band. All Landsat MSS bands were used in the classification, and the 60-m-resolution pixels were resampled to 30 m before classification by using a nearest neighbor approach.

Table 1. Landsat MSS and Landsat TM images used to develop Land-Use/Land-Cover maps.

Sensor	Image ID	Path	Row	Date
MSS	LM2038037007915310	38	37	1979-06-02
MSS	LM2038038007915310	38	38	1979-06-02
MSS	LM3038037007925210	38	37	1979-09-09
MSS	LM3038038007925210	38	38	1979-09-09
TM	L5036037_03719890609	36	37	1989-06-09
TM	L5036038_03819890609	36	38	1989-06-09
TM	L4035038_03819890610	35	38	1989-06-10
TM	L5035038_03819890821	35	38	1989-08-21
TM	L5036037_03719890929	36	37	1989-09-29
TM	L5036038_03819890929	36	38	1989-09-29
TM	L5035038_03819990614	35	38	1999-06-14
TM	L5036037_03719990621	36	37	1999-06-21
TM	L5036038_03819990621	36	38	1999-06-21
TM	L5035038_03819990817	35	38	1999-08-17
TM	L5036038_03819990824	36	38	1999-08-24
TM	L5036037_03719990824	36	37	1999-08-24
TM	L5035038_03820090508	35	38	2009-05-08
TM	L5036037_03720090616	36	37	2009-06-16
TM	L5036038_03820090616	36	38	2009-06-16
TM	L5036037_03720090904	36	37	2009-09-04
TM	L5036038_03820090904	36	38	2009-09-24
TM	L5035038_03820090929	35	38	2009-09-29

Atmospheric and Radiometric Calibration

Atmospheric effects were removed from the imagery using the cosine of theta (COST) model proposed by Chavez (1996). The model requires earth-sun distance and sun-elevation information for each collection date, the minimum digital number (DN) spectral radiance value for each band, and band bias and gain information (Chander and others, 2009). The images were corrected radiometrically from DNs to surface reflectance.

Image Transformations and Ancillary Data Sets

The following Landsat bands, image transformations, and ancillary datasets were used to develop the LULC classifications:

Landsat MSS (all data layers sampled to 30-m spatial resolution)

1. Slope (derived from USGS National Elevation Data 30-m DEMs).
2. Streams (rasterized buffer of streams polyline dataset).
3. 8 bands Landsat MSS in digital numbers (two seasons).
4. 8 bands multitemporal tasseled cap transformation (see appendix for MSS transformation values).

Landsat TM (all data layers in 30-m spatial resolution)

1. Slope (derived from USGS National Elevation Data 30-m DEMs).
2. Streams (rasterized buffer of streams polyline dataset).
3. Band 3 (red) 3×3 variance.
4. 12 bands Landsat TM reflectance (2 seasons).
5. 12 bands multitemporal Kauth Thomas (MKT) transformation.

The multitemporal tasseled cap and multitemporal Kauth Thomas (MKT) transformations (Collins and Woodcock, 1996) were calculated for multidate MSS and TM imagery. Multitemporal Kauth Thomas transformation is a linear change-detection technique similar to a multidate principal components analysis (PCA), where land-cover and phenological changes occurring between dates typically are manifested in one or more components (transformed data bands), providing useful land-cover signatures for the classification procedure (Collins and Woodcock, 1996). The advantage of using MKT over PCA is that MKT applies coefficients to the data, allowing for the comparison of results for multiscene data sets. We applied MKT to Landsat TM images using the coefficients described by Collins and Woodcock (1996). A Multitemporal tasseled cap matrix for the 4-band MSS data was created based on the single-date tasseled cap coefficients as described in Kauth and Thomas (1976). The multidate (8 band) matrix contains the first four standard tasseled cap coefficients and four additional orthogonal change coefficients. The eight tasseled cap coefficients were then normalized following the methods of Collins and Woodcock (1996) (appendix). It is important to note the Collins and Woodcock (1996) MKT coefficients for Landsat TM are for reflectance values, and the multitemporal tasseled cap coefficients for MSS are for DNs.

Training Data

Because of the size of the SCW (9,146 km^2), wall-to-wall high-resolution aerial photography/digital orthophoto quarter quadrangles (DOQQs) are too cumbersome to train and validate land cover for the four time periods; therefore, we trained the classifier and validated the data using 67 randomly sampled quadrangles in the U.S. portion of the watershed (of 243 possible quadrangles), and an additional seven quad-sized areas in Mexico where reference imagery was available (fig. 2). Training samples were identified as areas that did not display change from 1979 through 2009, allowing us to use a majority of the same samples for each time period. Polygons of homogeneous land-use and land-cover classes for each classification date were identified from aerial photographs and digital multispectral imagery and converted to raster for use with the Erdas Imagine “NLCD sampling tool” that is available as part of the “NLCD Mapping Tool” (tool download: http://www.mrlc.gov/).

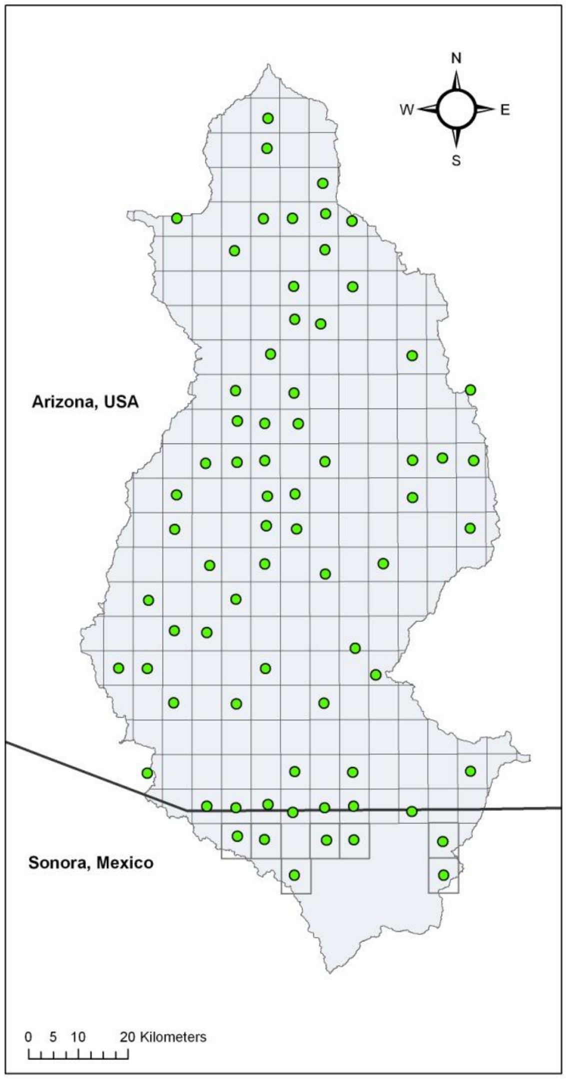


Figure 2. Random sample (green dots) of U.S. Geological Survey quadrangles used to train and assess the accuracy of land-use and land-cover change from 1979 through 2009 in the upper Santa Cruz watershed in southeastern Arizona, United States, and northern Sonora, Mexico.

Land-Cover Class Descriptions (From NLCD 2001)

We used the following descriptions from the NLCD 2001 classification scheme in our model for the SCW. A few noteworthy changes were made by us where class descriptions were modified to better describe Sonoran Desert land-cover types and arid riverine vegetation.

11. Open Water—All areas of open water, generally with less than 25 percent cover of vegetation or soil.

21. Developed, Open Space—Includes areas with a mixture of some constructed materials, but mostly vegetation in the form of lawn grasses. Impervious surfaces account for less than 20 percent of total cover. These areas most commonly include large-lot single-family housing units, parks, golf courses, and vegetation planted in developed settings for recreation, erosion control, or aesthetic purposes.

22. Developed, Low Intensity—Includes areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 20-49 percent of total cover. These areas most commonly include single-family housing units.

23. Developed, Medium Intensity—Includes areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 50-79 percent of the total cover. These areas most commonly include single-family housing units.

24. Developed, High Intensity—Includes highly developed areas where people reside or work in high numbers. Examples include apartment complexes, row houses and commercial/industrial. Impervious surfaces account for 80 to100 percent of the total cover.

31. Barren Land (Rock/Sand/Clay)—Barren areas of bedrock, desert pavement, scarps, talus, slides, volcanic material, glacial debris, sand dunes, strip mines, gravel pits, and other accumulations of earthen material. Generally, vegetation accounts for less than 15 percent of total cover.

41. Deciduous Forest—Areas dominated by trees generally greater than 2 m tall, and greater than 20 percent of total vegetation cover. More than 75 percent of the tree species shed foliage simultaneously in response to seasonal change.[1]

42. Evergreen Forest—Areas dominated by trees generally greater than 5 m tall, and greater than 20 percent of total vegetation cover. More than 75 percent of the tree species maintain their leaves all year. Canopy is never without green foliage.

52. Shrub/Scrub—Areas dominated by shrubs less than 5 m tall with shrub canopy typically greater than 20 percent of total vegetation. This class includes true shrubs, young trees in an early successional stage or trees stunted from environmental conditions.

71. Grassland/Herbaceous—Areas dominated by graminoid or herbaceous vegetation, generally greater than 80 percent of total vegetation. These areas are not subject to intensive management such as tilling, but can be utilized for grazing.

81. Pasture/Hay—Areas of grasses, legumes, or grass-legume mixtures planted for livestock grazing or the production of seed or hay crops, typically on a perennial cycle. Pasture/hay vegetation accounts for greater than 20 percent of total vegetation.[2]

82. Cultivated Crops—Areas used for the production of annual crops, such as corn, soybeans, vegetables, tobacco, and cotton, and also perennial woody crops, such as orchards and vineyards. Crop vegetation accounts for greater than 20 percent of total vegetation. This class also includes all land being actively tilled.

91. Palustrine Forested Wetlands—Includes all tidal and nontidal wetlands dominated by woody vegetation greater than or equal to 5 m in height, and all such wetlands that occur in tidal areas in which salinity due to ocean-derived salts is below 0.5 percent. Total vegetation coverage is greater than 20 percent.[3]

95. Emergent Herbaceous Wetlands - Areas where perennial herbaceous vegetation accounts for greater than 80 percent of vegetative cover, and the soil or substrate is periodically saturated with or covered with water.[4]

---

[1]In this dataset, cover class 41 is different than the Eastern Deciduous Forest type that is used in the NLCD. The Sonoran Desert is not home to broad-leafed forests, so we took the liberty of using class 41 to map the mesquite (*Prosopis* spp.) forests found along floodplains and xeric washes, which are deciduous and are an important component of the Sonoran desert ecosystem. Note tree height was changed from 5 m to 2 m.

[2]Includes class 85 (Urban/Recreational Grasses) from 1992 classification scheme.

[3]This class describes riparian forests and woodlands.

[4]This class describes riparian grasslands and marshes.

Classification and Regression Tree Model

We used C5 software (Quinlan, 1996) to develop the CART models and the NLCD Mapping Tool to classify the spatial data in Erdas Imagine. The CART model uses spectral and ancillary data as predictor variables and LULC classes as the response variables to create a dichotomous “tree” by recursively partitioning the training data into suitable class categories. Using the NLCD Mapping Tool, the classification tree rules derived from the training data are then applied directly to predictor variable layers to create a classified image. The classification process was iterative and involved some trial and error by (1) modifying portions of the training dataset that appeared to contribute to misclassification or class confusion, or (2) adding additional training samples if a particular class was “underperforming”.

Post-Processing

Upon completion of LULC modeling for each time period, we clipped the edges of all three classified Landsat scenes and mosaicked the images into one composite classification. We examined the classified products to locate any incongruous patches or areas with obviously misclassified pixels, particularly areas where the original images had clouds and cloud shadows or scanner errors (fig. 3). In these areas we used the ArcGIS “Raster Edit” extension to manually recode the areas with errors. Images were then processed using a majority filter with a 3×3

moving window, a process that eliminated most "noise" and misclassified pixels and gave the imagery a smoother and more generalized appearance. Once filtered, the data were clipped to the study area with the upper SCW boundary.

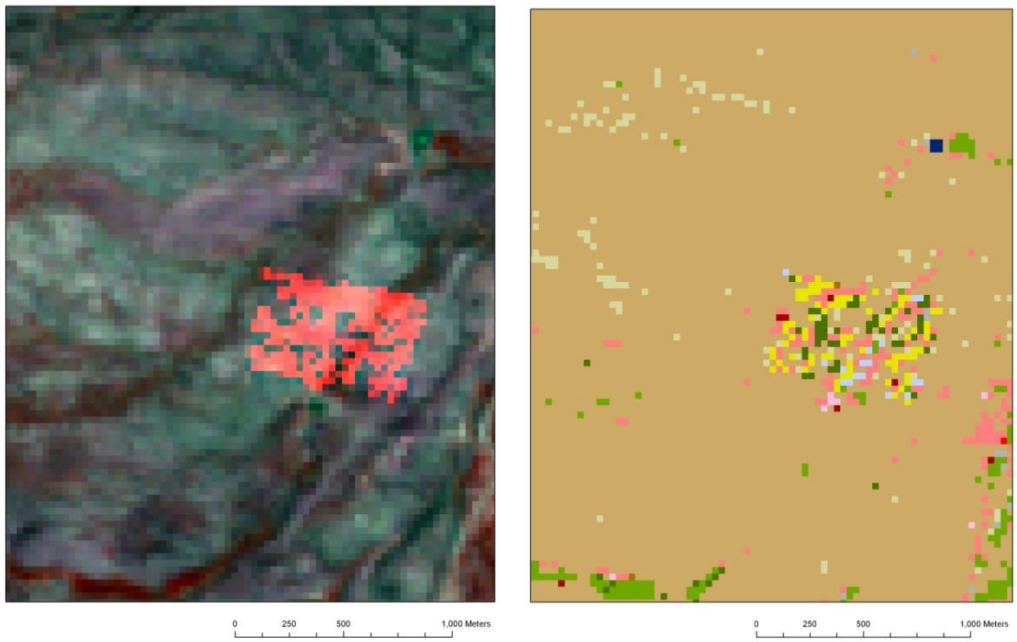


Figure 3. Example of data issues with original Landsat scene (left) and related classification errors (right). These errors were corrected manually by recoding the raster with the correct land-cover types, as determined from aerial photography.

Accuracy Assessment

We assessed the accuracy of the classified LULC data by determining the land-cover type from high-resolution aerial photographs acquired as near to the dates of the original Landsat images as possible. This type of accuracy assessment, while not as ideal as ground-based verification, is necessary when validating historical time series (Skirvin and others, 2004). It is important that the operator/interpreter doing the assessment has experience with aerial photo interpretation. M. Villarreal performed the accuracy assessment, given his training in photogrammetry and more than a decade of experience interpreting and mapping land cover from aerial photographs of the Sonoran Desert region.

A random stratified sample design was applied to a subset of the data (land-cover data were clipped on the basis of the total extent of the 67 randomly sampled quadrangles used for the training). We generated a thematic vector (polygon) layer from the raster data, clipped the vector data to the quadrangle's sample frame, and generated 50 random points stratified by the class-type polygons for a total of 700 validation points. This process was completed independently for each decadal land-cover map.

Each validation point was assigned a class value by the interpreter based on the aerial photography. Points were discarded if a class value could not be determined with certainty from the reference imagery. For example, older aerial photographs with vignetting around the outer edges of the frame made it difficult to assess the land-cover type; or, if the photo dates did not match perfectly with the Landsat acquisition, there may have been a potential land conversion in the interim. We removed class 11 (“Open Water”) from the accuracy assessment because the many small water bodies in the study area are seasonally ephemeral, making it difficult to assess accuracy with point-in-time aerial photographs (which are usually collected in the dry season).

To our knowledge there were no major aerial photo collections for 1979, so we assessed accuracy from multidate aerial collections: 1975, 1980, and 1983. Similarly, for 1999 there were no coincident DOQQs collected, so we used various datasets, primarily USGS high-resolution orthoimages of the Tucson Basin collected in 1998, 2000, and 2002 for the Pima Association of Governments and 1996 color-infrared DOQQs.

Results

The graphical and numerical output from each year’s LULC classification is presented in figures 4, 5, 6, and 7 and table 2.

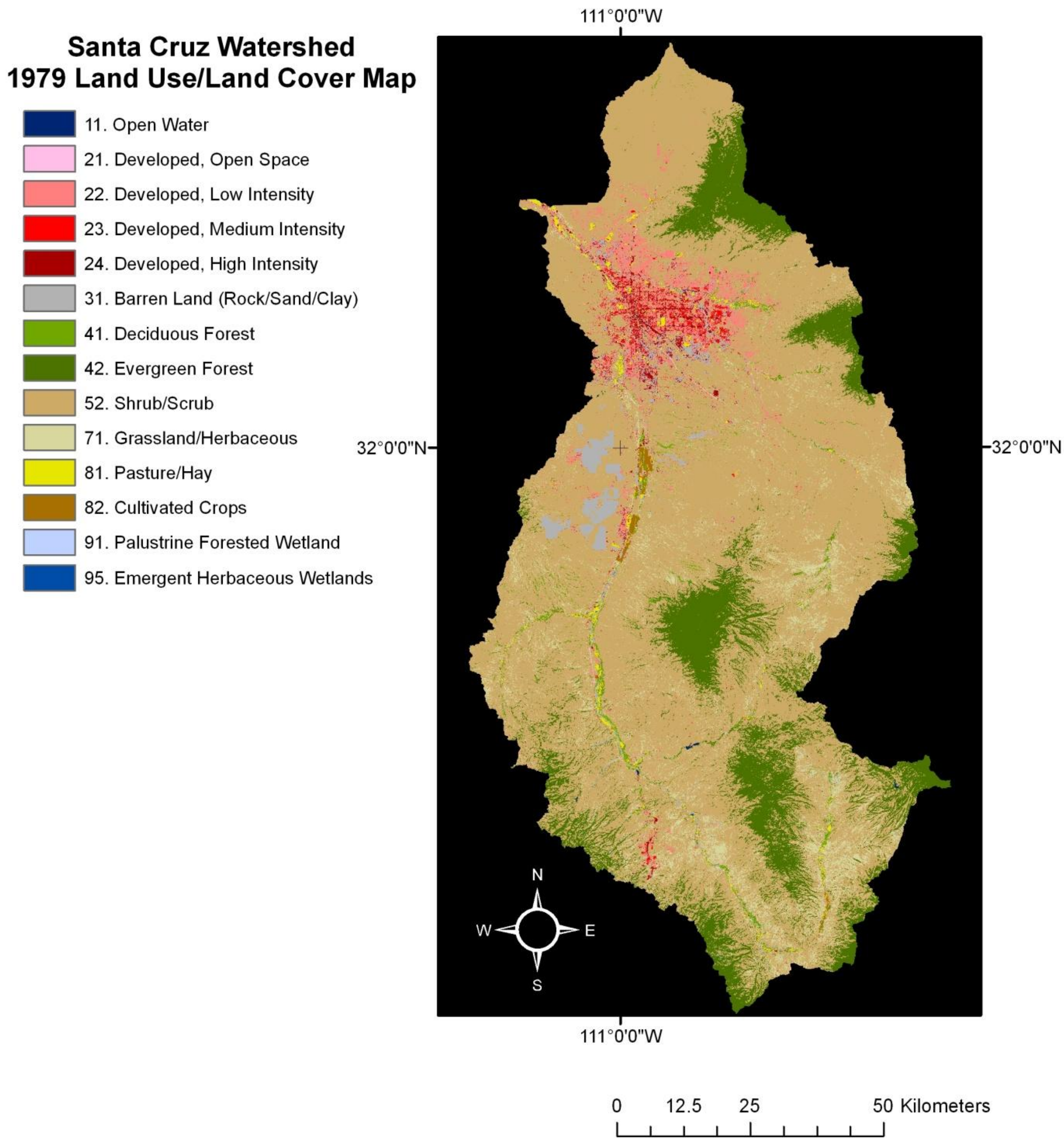


Figure 4. 1979 land-use/land-cover map of the Santa Cruz watershed.

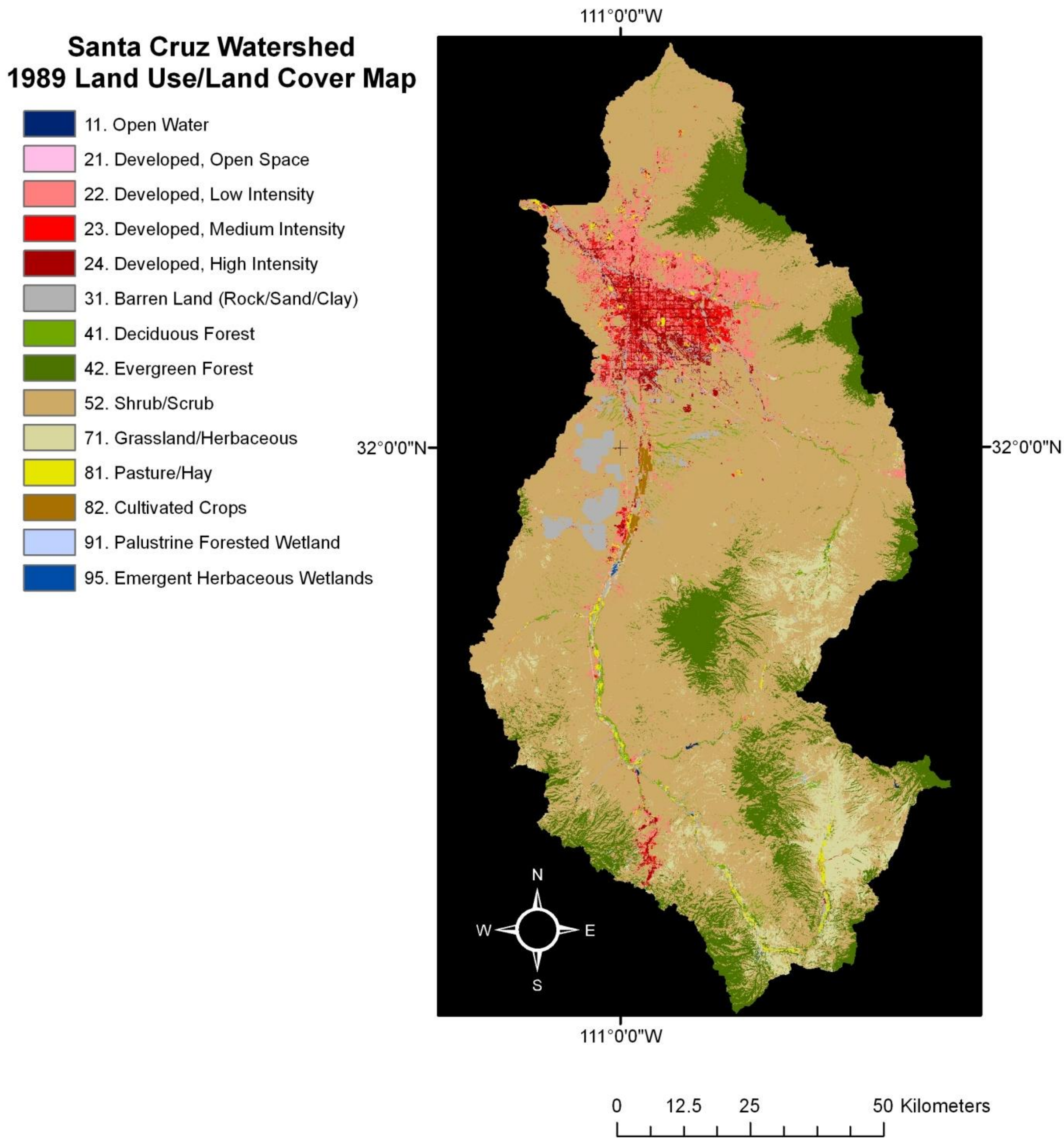


Figure 5. 1989 land-use/land-cover map of the Santa Cruz watershed.

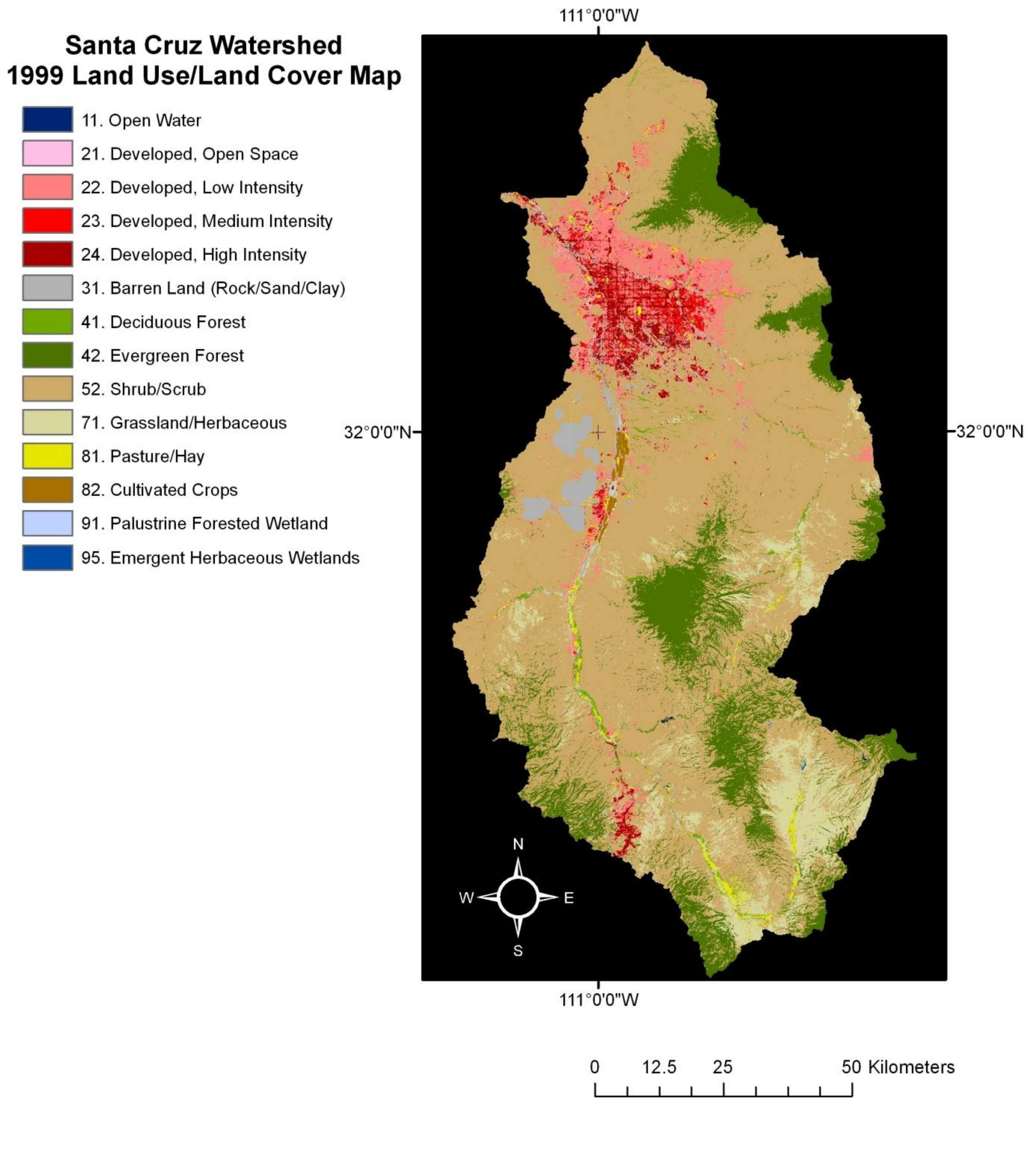


Figure 6. 1999 land-use/land-cover map of the Santa Cruz watershed.

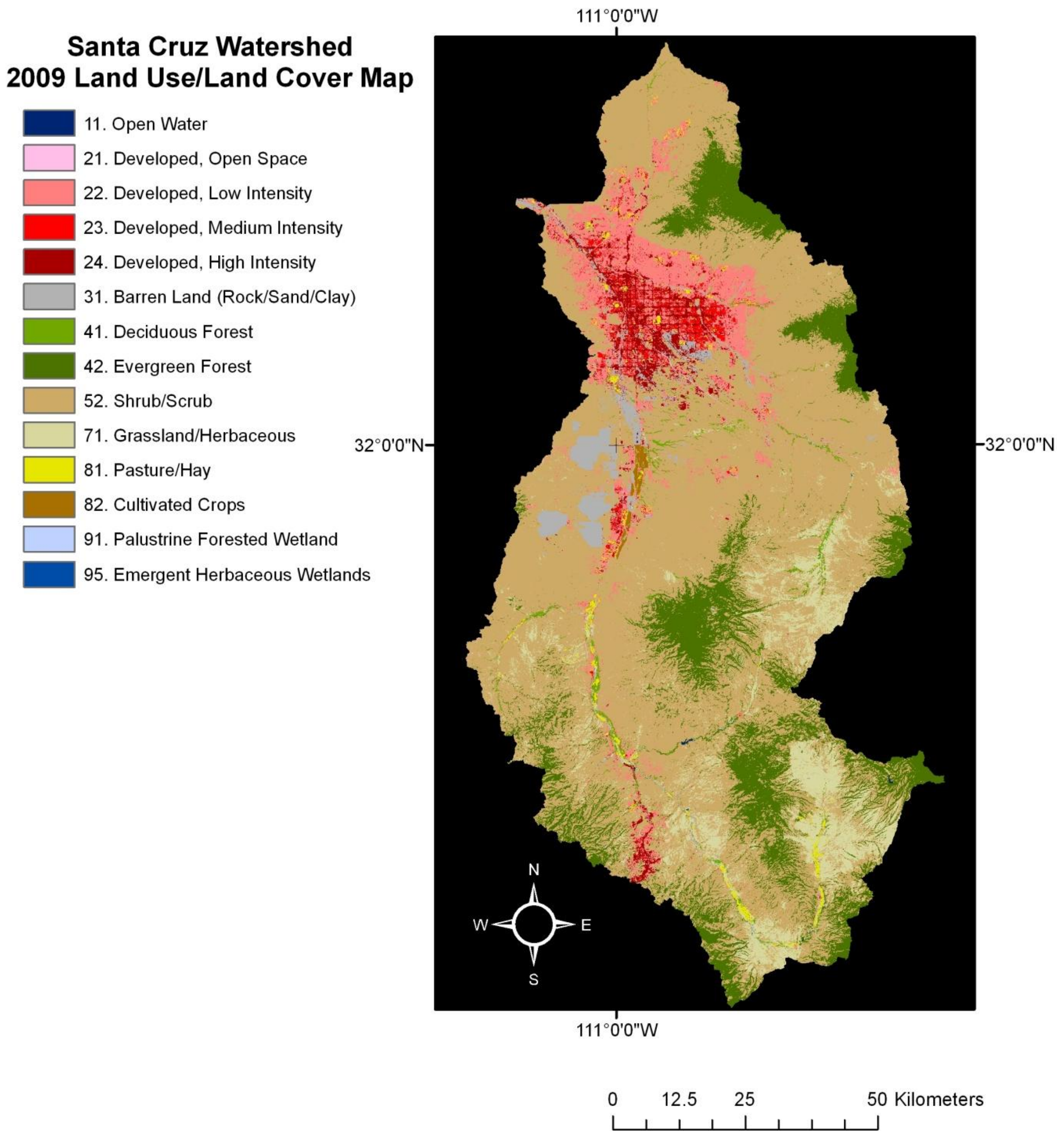


Figure 7. 2009 land-use/land-cover map of the Santa Cruz watershed.

Table 2. Class areas in the Santa Cruz watershed 1979-2009 (hectares).

Class description	Class value	Explanation	1979	1989	1999	2009
Water	**11**		339	284	282	315
Developed, Open Space	**21**		2,105	1,240	2,258	1,596
Developed, Low Intensity	**22**		32,130	47,184	59,458	68,044
Developed, Medium Intensity	**23**		6,644	9,351	8,642	9,687
Developed, High Intensity	**24**		8,237	19,460	21,418	21,877
Barren Land	**31**		23,131	23,128	21,618	23,857
Deciduous Forest	**41**		10,922	11,354	10,547	9,159
Evergreen Forest	**42**		132,709	120,667	129,467	131,964
Shrub/Scrub	**52**		624,739	609,755	577,926	563,778
Grassland/Herbaceous	**71**		63,484	63,038	72,190	74,741
Pasture/Hay	**81**		7,120	6,665	8,507	7,306
Cultivated Crops	**82**		2,803	2,365	2,380	2,545
Palustrine Forested Wetland	**91**		1,335	1,163	953	866
Emergent Herbaceous Wetlands	**95**		88	133	141	54

Accuracy Results

We measured the overall thematic accuracy of each map using the kappa statistic (Congalton, 1991). Kappa is a measure of agreement between map classification data and reference data, and its values range from –1.0 (complete disagreement) to 1.0 (perfect agreement). Overall accuracy, user's and producer's accuracy for each class, and a kappa statistic were derived from an error matrix (tables 3-6). In the following sections we will discuss some general trends in the accuracy statistics for all classes and focus on specific classes from each year with lower than average accuracy.

The LULC maps of 1979, 1989, 1999, and 2009 are more than 80 percent accurate overall (tables 3-6). Classes 23, 31, 41, 42, 52, 82 and 95 achieved >80 percent individual class accuracies for each classification date. Other class accuracies vary between dates to different degrees. For example, class 23 (Developed, Medium Intensity) had an accuracy of 83 percent in 1979 and 92 percent in each subsequent year, while class 21 (Developed, Open Space) displayed considerable overall variability (1979, 94 percent; 1989, 92 percent; 1999, 69 percent; and 2009, 78 percent).

The 1979 LULC map generally is accurate (overall accuracy, 84.72 percent; kappa, 0.83). Class 22 (Developed, Low Intensity, user accuracy, 72 percent) displayed some confusion with class 52 (Shrub/Scrub), which is understandable given that class 22 comprises a mixture of constructed materials and vegetation. Class 24 (Developed, High Intensity, user accuracy, 68 percent) shows some confusion with other Developed classes and Barren Land. Class 91 (Palustrine Forested Wetland, user accuracy, 71 percent) was confused with a number of other land cover classes, but primarily with class 52 (Shrub/Scrub). It is important to note that in 1979, because the total area of class 91was small, a large number of random points usually were distributed in single misclassified pixels located in a cultivated area along the floodplain.

The 1989 LULC map generally is accurate (overall accuracy, 84.30 percent; kappa, 0.83). As in the 1979 map, class 22 (Developed, Low Intensity, user accuracy, 69 percent) was primarily confused with class 52 (Shrub/Scrub). Class 91 (Palustrine Forested Wetland) displayed a relatively low class accuracy (user accuracy, 57 percent) and was confused with a number of other land cover classes, but primarily with class 71 (Grassland/Herbaceous).

The 1999 LULC map has an overall accuracy of 81.79 percent and kappa of 0.80, making it the least accurate of the LULC datasets. It is unclear why map accuracy was lower for this particular year, but it may be related to the fact that reference photographs were from several different dates, and none were acquired in 1999 proper (see above section). As in the 1979 and 1989 maps, class 22 (Developed, Low Intensity, user accuracy, 66 percent) primarily was confused with class 52 (Shrub/Scrub), as well as class 21 (Developed, Open Space, user accuracy, 69 percent). Class 91 (Palustrine Forested Wetland) displayed a relatively low accuracy (user accuracy, 60 percent) and was (mis)classified as a number of other land-cover classes.

The 2009 LULC map achieved the highest accuracy of the four (overall accuracy, 86.50 percent; kappa, 0.85). This high accuracy may be attributed to the high quality of the reference dataset—seamless, high-resolution NAIP imagery collected in 2010.

Like the caveats mentioned above concerning the low accuracy of some classes with a small proportion of the total cover, it also may be the case that the accuracy of some small-area classes was erroneously high. For example, in 2009, of the 50 randomly generated points for class 95 (Emergent Herbaceous Wetlands), 45 fell within two large land-cover patches, one of which was used to train the classifier. The resulting accuracy for that class was 100 percent. However, the large patch that was not used in the training is the wetland at Las Cienegas, indicating the successful classification of one of the rare wetlands in the study area.

Conclusions

This dataset includes four dates of accurate, moderate-resolution (30 m) land-use/land-cover (LULC) maps for the upper portion of the binational Santa Cruz watershed. The results of our accuracy assessments indicate that the LULC maps are above 80 percent accurate and will provide realistic input data layers for future watershed models. Individual class accuracies vary by year and by cover class, and individual error matrices should be consulted when focusing analyses on changes in specific LULC types.

Table 3. Error matrix for 1979 land-use/land-cover map.

		21	22	23	24	31	41	42	52	71	81	82	91	95	Total	User	Commission	Kappa
Developed, Open Space	**21**	**45**	0	0	0	0	0	0	3	0	0	0	0	0	48	94%	6%	0.9317
Developed, Low Intensity	**22**	1	**36**	2	0	0	3	0	7	1	0	0	0	0	50	72%	28%	0.7020
Developed, Medium Intensity	**23**	2	0	**40**	1	1	0	0	2	2	0	0	0	0	48	83%	17%	0.8202
Developed, High Intensity	**24**	2	1	3	**34**	4	1	0	3	0	2	0	0	0	50	68%	32%	0.6595
Barren Land	**31**	0	0	0	2	**43**	0	0	3	0	1	0	0	0	49	88%	12%	0.8665
Deciduous Forest	**41**	0	0	0	0	0	**44**	0	4	0	1	0	0	0	49	90%	10%	0.8885
Evergreen Forest	**42**	0	0	0	0	0	1	**48**	1	0	0	0	0	0	50	96%	4%	0.9565
Shrub/Scrub	**52**	1	0	0	0	0	0	0	**49**	0	0	0	0	0	50	98%	2%	0.9767
Grassland/Herbaceous	**71**	1	0	0	0	0	0	1	8	**39**	1	0	0	0	50	78%	22%	0.7630
Pasture/Hay	**81**	0	0	0	0	3	0	1	1	1	**41**	0	1	1	49	84%	16%	0.8226
Cultivated Crops	**82**	0	0	0	0	0	1	0	0	1	2	**45**	0	0	49	92%	8%	0.9110
Palustrine Forested Wetland	**91**	0	0	0	0	0	2	0	5	0	1	2	**27**	1	38	71%	29%	0.6967
Emergent Herbaceous Wetlands	**95**	0	0	0	0	0	0	0	0	0	0	4	0	**30**	34	88%	12%	0.8759
	Total	52	37	45	37	51	52	50	86	44	49	51	28	32	**614**			
	Producer	87%	97%	89%	92%	84%	85%	96%	57%	89%	84%	88%	96%	93.75%		**521**		
	Omission	13%	3%	11%	8%	16%	15%	4%	43%	11%	16%	12%	4%	6.25%			**84.85%**	
	Kappa	85%	97%	88%	91%	83%	83%	96%	53%	88%	82%	87%	96%	0.93				**0.8357**

Table 4. Error matrix for 1989 land-use/land-cover map.

		21	22	23	24	31	41	42	52	71	81	82	91	95	Total	User	Commission	Kappa
Developed, Open Space	**21**	**44**	2	1	0	0	0	0	1	0	0	0	0	0	48	92%	8%	0.9096
Developed, Low Intensity	**22**	3	**33**	2	1	0	1	2	5	1	0	0	0	0	48	69%	31%	0.6670
Developed, Medium Intensity	**23**	0	3	**45**	0	1	0	0	0	0	0	0	0	0	49	92%	8%	0.9110
Developed, High Intensity	**24**	0	0	3	**38**	3	1	0	0	3	0	0	0	0	48	79%	21%	0.7776
Barren Land	**31**	1	0	0	0	**40**	0	0	3	0	0	0	0	0	44	91%	9%	0.9021
Deciduous Forest	**41**	0	0	0	0	0	**44**	1	2	0	0	1	0	0	48	92%	8%	0.9085
Evergreen Forest	**42**	0	0	0	0	0	0	**46**	3	1	0	0	0	0	50	92%	8%	0.9131
Shrub/Scrub	**52**	0	0	0	0	0	0	0	**47**	2	0	0	0	0	49	96%	4%	0.9530
Grassland/Herbaceous	**71**	0	0	0	0	0	1	0	10	**34**	0	0	0	1	46	74%	26%	0.7152
Pasture/Hay	**81**	0	0	0	0	0	2	0	1	3	**35**	0	1	0	42	83%	17%	0.8221
Cultivated Crops	**82**	0	0	0	0	0	1	0	0	0	0	**46**	1	0	48	96%	4%	0.9547
Palustrine Forested Wetland	**91**	0	0	0	0	0	5	0	4	8	4	0	**28**	0	49	57%	43%	0.5496
Emergent Herbaceous Wetlands	**95**	0	0	0	0	0	0	0	5	0	0	3	0	**41**	49	84%	16%	0.8248
	Total	48	38	51	39	44	55	49	81	52	39	50	30	42	**618**			
	Producer	92%	87%	88%	97%	91%	80%	94%	58%	65%	90%	92%	93%	97.62%		**521**		
	Omission	8%	13%	12%	3%	9%	20%	6%	42%	35%	10%	8%	7%	2.38%			**84.30%**	
	Kappa	91%	86%	87%	97%	90%	78%	93%	54%	63%	89%	91%	93%	0.97				**0.8299**

Table 5. Error matrix for 1999 land-use/land-cover map.

		21	22	23	24	31	41	42	52	71	81	82	91	95	Total	User	Commission	Kappa
Developed, Open Space	**21**	**34**	3	0	0	4	0	0	8	0	0	0	0	0	49	69%	31%	0.6739
Developed, Low Intensity	**22**	2	**35**	2	0	5	3	0	6	0	0	0	0	0	53	66%	34%	0.6364
Developed, Medium Intensity	**23**	0	4	**46**	0	0	0	0	0	0	0	0	0	0	50	92%	8%	0.9127
Developed, High Intensity	**24**	0	0	4	**43**	2	0	0	1	0	0	0	0	0	50	86%	14%	0.8499
Barren Land	**31**	2	0	0	0	**45**	0	0	2	1	0	0	0	0	50	90%	10%	0.8898
Deciduous Forest	**41**	0	0	0	0	1	**39**	0	3	0	0	2	2	3	50	78%	22%	0.7621
Evergreen Forest	**42**	0	0	0	0	0	1	**43**	6	0	0	0	0	0	50	86%	14%	0.8496
Shrub/Scrub	**52**	1	0	1	0	0	0	1	**47**	0	0	0	0	0	50	94%	6%	0.9304
Grassland/Herbaceous	**71**	0	0	0	0	0	0	0	5	**42**	1	0	0	1	49	86%	14%	0.8452
Pasture/Hay	**81**	0	0	0	0	0	2	0	1	3	**39**	3	0	2	50	78%	22%	0.7625
Cultivated Crops	**82**	0	0	0	0	0	1	0	1	0	3	**43**	0	0	48	90%	10%	0.8872
Palustrine Forested Wetland	**91**	0	0	0	0	2	2	0	5	0	4	1	**25**	3	42	60%	40%	0.5773
Emergent Herbaceous Wetlands	**95**	0	0	0	0	0	0	0	3	3	0	0	0	**40**	46	87%	13%	0.8587
	Total	39	42	53	43	59	48	44	88	49	47	49	27	49	**637**			
	Producer	87%	83%	87%	100%	76%	81%	98%	53%	86%	83%	88%	93%	81.63%		**521**		
	Omission	13%	17%	13%	0%	24%	19%	2%	47%	14%	17%	12%	7%	18.37%			**81.79%**	
	Kappa	86%	82%	86%	100%	74%	80%	98%	49%	85%	82%	87%	92%	0.80				**0.8026**

Table 6. Error matrix for 2009 land-use/land-cover map.

		21	22	23	24	31	41	42	52	71	81	82	91	95	Total	User	Commission	Kappa
Developed, Open Space	21	39	1	1	3	2	0	0	3	0	1	0	0	0	50	78%	22%	0.7637
Developed, Low Intensity	22	2	40	3	0	1	0	0	3	0	0	0	0	0	49	82%	18%	0.8020
Developed, Medium Intensity	23	0	3	46	1	0	0	0	0	0	0	0	0	0	50	92%	8%	0.9119
Developed, High Intensity	24	2	0	6	37	5	0	0	0	0	0	0	0	0	50	74%	26%	0.7217
Barren Land	31	0	0	1	0	46	0	0	2	1	0	0	0	0	50	92%	8%	0.9124
Deciduous Forest	41	0	0	0	0	0	42	0	4	2	0	0	1	0	49	86%	14%	0.8447
Evergreen Forest	42	0	0	0	0	0	2	44	4	0	0	0	0	0	50	88%	12%	0.8702
Shrub/Scrub	52	0	1	0	0	0	0	0	47	0	0	0	0	0	48	98%	2%	0.9763
Grassland/Herbaceous	71	0	0	0	0	0	0	0	8	39	0	0	0	0	47	83%	17%	0.8152
Pasture/Hay	81	0	0	0	0	0	0	0	1	6	39	0	0	0	46	85%	15%	0.8374
Cultivated Crops	82	0	0	0	0	0	1	0	1	0	0	47	0	0	49	96%	4%	0.9558
Palustrine Forested Wetland	91	0	0	0	0	0	5	3	2	1	0	1	23	0	35	66%	34%	0.6434
Emergent Herbaceous Wetlands	95	0	0	0	0	0	0	0	0	0	0	0	0	49	49	100%	0%	1.0000
	Total	43	45	57	41	54	50	47	75	49	40	48	24	49	622			
	Producer	91%	89%	81%	90%	85%	84%	94%	63%	80%	98%	98%	96%	100.00%		538		
	Omission	9%	11%	19%	10%	15%	16%	6%	37%	20%	3%	2%	4%	0.00%			86.50%	
	Kappa	90%	88%	79%	89%	84%	83%	93%	60%	78%	97%	98%	96%	1.00				0.8536

Acknowledgments

The authors wish to thank Terence Arundel, Sam Drake, and Willem van Leeuwen for their reviews of this material. The authors also would like to thank Dr. Stuart Marsh of the Arizona Remote Sensing Center (ARSC) for hardware and software support and Kyle Hartfield of ARSC for GIS assistance.

References Cited

Chander, G., Markham, B.L., and Helder, D.L., 2009, Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors: Remote Sensing of Environment, v. 113, p. 893-903.

Chavez, P.S., Jr., 1996, Image-based atmospheric corrections—revisited and improved: Photogrammetric Engineering and Remote Sensing, v. 62, no. 9, p. 1025-1036.

Collins, J.B., and Woodcock, C.E., 1996, An assessment of several linear change detection techniques for mapping forest mortality using multi-temporal Landsat TM data: Remote Sensing of Environment, v. 56, p. 66–77.

Congalton, R.G., 1991, A review of assessing the accuracy of classifications of remotely sensed data: Remote Sensing of Environment, v. 37, p. 35-46.

Homer, C., Dewitz, J., Fry, J., Coan, M., Hossain, N., Larson, C., Herold, N., McKerrow, A., VanDriel, J.N., and Wickham J., 2007, Completion of the 2001 National Land Cover Database for the Conterminous United States: Photogrammetric Engineering and Remote Sensing, v. 73, no. 4, p. 337-341.

Kauth, R.J., and Thomas, G.S., 1976, The tasseled cap—a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat, *in* Proceedings of the symposium on machine processing of remotely sensed data: West Lafayette, Indiana, Purdue University, p. 4B41-4B51.

Norman, L., Tallent-Halsell, N., Labiosa, W., Weber, M., McCoy, A., Hirschboeck, K., Callegary, J., van Riper, C., III, and Gray, F., 2010, Developing an ecosystem services online decision support tool to assess the impacts of climate change and urban growth in the Santa Cruz Watershed—where we live, work, and play: Sustainability v. 2, p. 2044-2069.

Quinlan, J.R., 1996, Bagging, boosting and C4.5, *in* Proceedings of the 13th national conference on artificial intelligence: Portland, Oregon, AAAI Press, p. 725–730.

Riitters, K.H., and Wickham, J.D., 2003, How far to the nearest road?: Frontiers in Ecology and the Environment, v. 1, p. 125-129.

Scanlon, B.R., Reedy, R.C., Stonestrom, D.A., Prudic, D.E., and Dennehy, K.F., 2005, Impact of land use and land cover change on groundwater recharge and quality in the southwestern US: Global Change Biology, v. 11, p. 1577-1593.

Schulte, L., Mladenoff, D., Crow, T., Merrick, L., and Cleland, D., 2007, Homogenization of northern U.S. Great Lakes forests due to land use: Landscape Ecology, v. 22, p. 1089-1103.

Skirvin, S.M., Kepner, W.G., Marsh, S.E., Drake, S.E., Maingi, J.K., Edmonds, C.M., Watts, C.J., and Williams, D.R., 2004, Assessing the accuracy of satellite-derived land-cover classification using historical aerial photography, digital orthophoto quadrangles, and airborne video data, *in* Lunetta, R.S., and Lyon, J.G., eds., Remote sensing and GIS accuracy assessment: Boca Raton, Florida, CRC Press.

Vogelmann, J.E., Howard, S.M., Yang, L., Larson, C.R., Wylie, B.K., and Van Driel, J.N., 2001, Completion of the 1990's National Land Cover Data Set for the conterminous United States: Photogrammetric Engineering and Remote Sensing, v. 67, p. 650-662.

Appendix. Multitemporal Tasseled Cap Coefficients for Landsat MSS Data

Original							
Brightness	Greenness	Wetness	Non-such	ΔBrightness	ΔGreenness	ΔWetness	ΔNon-such
0.433	-0.29	-0.829	0.223	0.433	0.632	0.586	0.264
0.632	-0.562	0.522	0.012	-0.29	-0.562	0.6	0.491
0.586	0.6	-0.039	-0.543	-0.829	0.522	-0.039	0.194
0.264	0.491	0.194	0.81	0.223	0.012	-0.543	0.81

Normalized							
Brightness	Greenness	Wetness	Non-	ΔBrightness	ΔGreenness	ΔWetness	ΔNon-such
0.3062	-0.2051	-0.5862	0.1577	-0.3062	0.2051	0.5862	-0.1577
0.4469	-0.3974	0.3691	0.0085	-0.4469	0.3974	-0.3691	-0.0085
0.4144	0.4243	-0.0276	-0.3840	-0.4144	-0.4243	0.0276	0.3840
0.1867	0.3472	0.1372	0.5728	-0.1867	-0.3472	-0.1372	-0.5728
0.3062	-0.2051	-0.5862	0.1577	0.3062	-0.2051	-0.5862	0.1577
0.4469	-0.3974	0.3691	0.0085	0.4469	-0.3974	0.3691	0.0085
0.4144	0.4243	-0.0276	-0.3840	0.4144	0.4243	-0.0276	-0.3840
0.1867	0.3472	0.1372	0.5728	0.1867	0.3472	0.1372	0.5728