INT. J. REMOTE SENSING, 20 DECEMBER, 2004,
VOL. 25, NO. 24, 5637–5654

Using AVIRIS data and multiple-masking techniques to map urban forest tree species

Q. XIAO*, S. L. USTIN

Department of Land, Air, and Water Resources, University of California, Davis, CA 95616, USA

and E. G. McPHERSON

USDA Forest Service, Pacific Southwest Research Station, Center for Urban Forest Research, Davis, CA 95616, USA

(*Received 15 September 2003; in final form 1 June 2004*)

Abstract. Tree type and species information are critical parameters for urban forest management, benefit cost analysis and urban planning. However, traditionally, these parameters have been derived based on limited field samples in urban forest management practice. In this study we used high-resolution Airborne Visible Infrared Imaging Spectrometer (AVIRIS) data and multiple-spectral masking techniques to identify and map urban forest trees. Trees were identified based on their spectral character difference in AVIRIS data. The use of multiple-masking techniques shift the focus to the target land cover types only, thus reducing confounding noise during spectral analysis. The results were checked against ground reference data and by comparison to tree information in an existing geographical information system (GIS) database. At the tree type level, mapping was accomplished with 94% accuracy. At the tree species level, the average accuracy was 70% but this varied with both tree type and species. Of the four evergreen tree species, the average accuracy was 69%. For the 12 deciduous tree species, the average accuracy was 70%. The relatively low accuracy for several deciduous species was due to small tree size and overlapping among tree crowns at the 3.5 m spatial resolution of AVIRIS data. This urban forest tree species mapping method has the potential to increase data update intervals and accuracy while reducing costs compared to field sampling or other traditional methods.

1. Introduction

Over 70% of the population within developed countries lives in cities. Worldwide, the average proportion of the urban population is 42% (World Population Reference 1993). Urbanization creates significant changes in land use and land cover, affecting the structure, pattern and function of the ecosystem (Douglas 1983). The public is increasingly concerned about how these changes influence daily life and affect the sustainability of 'quality of life' for future

*Corresponding author; Qingfu Xiao, CUFR, c/o Department of Environmental Horticulture, University of California, Davis, CA 95616, USA; e-mail: qxiao@ucdavis.edu

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2004 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/0143116041233129122 4

generations (World Resources Institute 1996). The structure and function of urban ecosystems can be studied using the same methods as the study of natural environments (Rowntree 1984).

Vegetation canopy is an important land cover type that affects the development of the urban ecosystem (McBride and Jacobs 1986). Trees planted in urban settings are described as an urban forest, and they are an important part of the urban ecosystem. Urban trees play important roles in improving landscape aesthetics, reducing pollution and moderating the urban energy budget, water use, reducing storm runoff, and providing other amenities (Dwyer *et al.* 1992, Xiao and McPherson 2003). The increase in the proportion of pavement area during the process of urbanization strongly influences energy exchange, hydrology and micro-climate (Arnold and Gibbons 1996). Many problems facing management of the urban ecosystem are related to these factors. For example, urban heat island effects and increased storm runoff are related to vegetation cover (USDA 1975, Gallo and Tarpley 1996). Air quality and water use are related to crown density and tree density because the total leaf surface area and leaf surface area per unit of land area controls both air pollutant removal and evapotranspiration rates (Peper and McPherson 2003). Understanding impacts on air quality, energy partitioning and hydrologic processes in the urban ecosystem depend on knowledge of tree species, leaf and stem surface areas, tree dimensions, and percentage of pavement cover, among other things. To understand how urban forests function and to estimate the value of their environmental services we must first recognize properties related to urban forest structure and composition (McPherson *et al.* 1997). Also, a good understanding of the structure of the urban forest provides other information useful to urban managers, such as for planning tree pruning, removal, and insect or disease control activities.

Basic information required to describe urban forest structure includes tree numbers, spatial distributions, species composition, dimensions and growing conditions. Traditionally, this information is collected in field surveys. However, such surveys are expensive and time consuming, and require periodic updates to remain valid. Aerial photograph interpretation has been used successfully but it is slow and expensive to conduct the mapping at large scale. Vegetation has unique spectral reflectance characteristics with strong absorption in red wavelengths and strong reflectance in near-infrared wavelengths, which allow separation of plants from other ground surface covers. Differences in the allocation of foliage, stems and varied architecture among tree species may provide sufficient information to uniquely identify them with Airborne Visible Infrared Imaging Spectrometer (AVIRIS) data, an airborne hyperspectral imaging instrument that can collect data at 3–4 m pixels. Differences in canopy architecture such as leaf area density, leaf and branch zenith angles, leaf shape, internal anatomy, and leaf and branch surface roughness cause different tree species to have different reflectance spectra. The NDVI (Normalized Difference Vegetation Index), red-edge, and other band ratio methods have been used for separating different amounts of vegetation. However, these simple methods cannot be used to identify tree species because they do not capture the unique features of the specific spectral characteristics of specific tree species. While the texture analysis method works well in natural forest mapping (Franklin *et al.* 2001) it does not work well in the urban forest because species composition and tree density are heterogeneous at high spatial resolutions. The spectral reflectance characteristics of different tree species differ with the relative proportion of biochemicals (e.g. photosynthetic pigments, cell wall materials, water

concentrations) and the scattering properties affected by the internal leaf structures. Retrieval of canopy biochemistry has been demonstrated several times as has the use of these techniques in mapping individual species (Underwood *et al.* 2003). At the canopy scale, the three dimensional structure and distribution of stems and leaves of the trees influence reflectance. New high spatial and spectral resolution remote sensing technology brings us an opportunity to abstract spatially explicit urban forest information from remote sensing data. Also, it provides a mechanism for tracking and monitoring tree health and canopy cover changes through repeated data acquisition. With several new hyperspatial and hyperspectral digital imaging systems available on aircraft and in space, e.g. Quickbird and Hyperion, respectively, we can investigate the use of high-resolution spectral data for characterizing and monitoring urban trees.

In principle, images recorded by airborne or satellite-based sensors can be obtained at reasonably frequent intervals, at desired spatial and spectral resolutions and at lower cost per unit land area compared with traditional field survey methods (Martin *et al.* 1988, Ehlers 1990). Remotely sensed data have been used for identifying and mapping vegetation, land use and land cover in many regional or sub-regional assessments (Morgan *et al.* 1993, Huang *et al.* 1995, Nowak *et al.* 1996, Huang and Ridd 2002, Price *et al.* 2002, Arthur-Hartranft *et al.* 2003, Clapham 2003, Huang and Townshend 2003, Weber and Puissant 2003). Strong correlations between Landsat Thematic Mapper (TM) data and crown closure have been found in rural forests where the trees species were known (Franklin *et al.* 2003, Xu *et al.* 2003). However, accuracy estimates in urban settings become a problem due to the complex spatial assemblages of disparate patches of land cover types. Urban areas are a mosaic of many tree types (e.g. species and dimensions), land uses, and man-made structures, each of which has different spectral reflectance characteristics (Gong and Howarth 1990). Unlike trees in rural forests, which tend to form continuous canopies, trees in urban settings are often single trees or isolated groups. The influence of background, such as soil and shadow, makes the problem of characterizing trees by remote sensing even more difficult. In such cases, high spatial resolution of remotely sensed data is important for mapping individual trees (Avery and Berlin 1992).

In boreal forests, Ustin and Xiao (2001) showed that even at the same spatial resolution vegetation mapping accuracy is significantly improved using the hyperspectral 20 m AVIRIS data compared to three-band 20 m Satellite pour l'Observation de la Terre (SPOT) data, especially for estimating dominant tree species. However, mapping species in an urban landscape presents many challenges not found in natural forest environments.

NASA's (National Aeronautics and Space Administration) AVIRIS instrument is a hyperspectral imaging system that delivers calibrated images of the upwelling spectral radiance in 224 contiguous spectral channels with wavelengths from 400 nm to 2500 nm (Green *et al.* 1998). Low altitude AVIRIS data were acquired at 3.5 m spatial resolution, giving us an opportunity to study urban forest at the single tree level. These enriched spatial and spectral data reduce the resolution problems associated with broad-band low-spatial resolution sensors.

Coupling geographical information systems (GIS) to the analysis of remote sensing data improved the accuracy of the results. Incorporation of spatial location has become a standard method for registering images to base maps, as shown in recent reports (e.g. Grimmond and Souch 1994, Blackburn and Milton 1997, Ambrosia *et al.* 1998, Lakshmi *et al.* 1998, Li 1998, Shao *et al.* 1998). Our ability to

accurately locate individual trees using the GIS database makes abstraction of the spectral reflectance characteristics from AVIRIS data relatively easy. In this study, we demonstrate an important application of urban forest characterization by combining remote sensing and GIS techniques.

2. Objectives

There were three objectives for this study. The first objective was to identify urban tree species by physiognomic type based on their spectral character as detected by the AVIRIS sensor, that is, whether they are broadleaf deciduous, broadleaf evergreen or conifer types. The second objective was to identify urban trees by species based on their canopy reflectance characteristics. The third objective was to map these urban trees. Each of these levels is useful for providing tree canopy information to urban planning and projects related to analysis of regional urban energy budgets, air pollution and hydrology.

3. Methodology

3.1. *Study site*

We selected the City of Modesto, California, USA (latitude: 37° 38′10″ N, longitude: 121° 11′10″ W) as our study site (figure 1). The city is located in the Central Valley of California and has a population of 186 000 with more than 70% of the population living in family households. City development began in the mid-1800s. Like many valley cities near the Sierra Nevada Mountains, it is undergoing a period of rapid population growth and expansion. Trees in Modesto are diverse in both species type and dimension. There are minimal topographic gradients in the study area or in the surrounding region. The study area elevation range was approximately 11 m.

3.2. *Datasets*

To investigate the use of imaging spectrometry for monitoring urban forests, three types of datasets were collected from the study area: a GIS-based street tree database (field survey of all street trees), high-resolution AVIRIS data, and GIS layers for locations such as streets and parcels.

The GIS database includes layers of base maps, parcels, street trees, soils, and

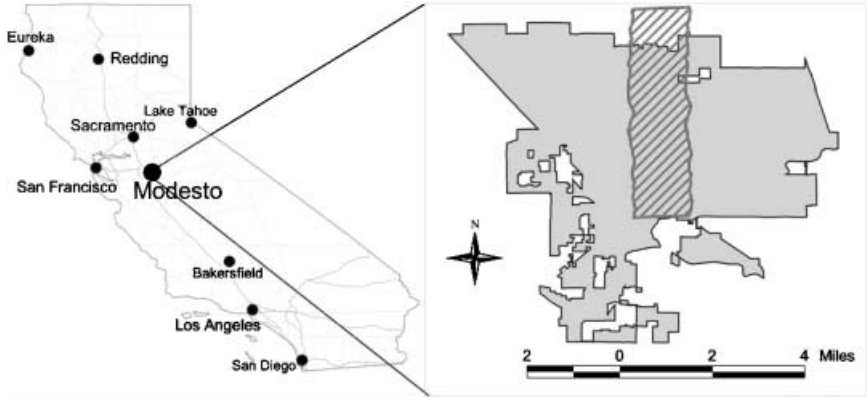


Figure 1. Study site. The shaded area is the Airborne Visible Infrared Imaging Spectrometer (AVIRIS) flight area.

land use. These data were provided by the Engineering and Transportation Department of the City of Modesto, California. The citywide street tree database contains 184 tree species and 75 629 individual trees. Most of these trees are broadleaf deciduous (87%) and the broadleaf evergreen and conifer tree types only account for 7% and 6%, respectively. Information for each tree includes: species code, scientific name and common name, tree ID number, year tree planted, and the access address (e.g. street address, city area, corner street and corner address). The tree layer in GIS was generated from tree survey spreadsheet and the trees street address and street GIS layer based on the address matching method (ArcView, Environmental Systems Research Institute Inc. 1997).

During the summer of 1998, 648 trees were measured in the city. The random sample consisted of approximately 30 trees from each of the 22 most common species (table 1). Trees belonging to these 22 species accounted for over 90% of the entire street tree population. Street address was used to locate tree samples. Field measurements included both tree dimension and maintenance information such as: DBH (diameter at breast height, or the diameter of the bole at roughly 1.5 m height), tree crown height, bole height at the bottom of the crown, crown diameter, total tree leaf surface area, geometric crown shape, site index, health/condition and tree pruning rating (Peper *et al.* 2001). Based on these field measurements, a

Table 1. Tree species of the Modesto study site and number of trees included within the AVIRIS data with tree crowns greater than one pixel.

Species					
Scientific name	Common name	Type	Code	Total sample	Crown diameter > 1
Acer saccharinum	Silver Maple	BD	ACSA	8	6
Betula pendulata	Birch	BD	BEPE	6	1
Celtis sinensis	Chinese Hackberry	BD	CESI	3	3
Cinnamomum camphora	Camphor	BE	CICA	9	7
Fraxinus excelsior 'Hessii'	Hess Ash	BD	FREX	2	2
Fraxinus × Moraine	Moraine Ash	BD	FRMO	5	5
Fraxinus oxycarpa 'Raywood'	Raywood Ash	BD	FROX	3	3
Fraxinus pennsylvanica 'Marshall'	Marshal Ash	BD	FRPE	11	10
Fraxinus velutina 'Modesto'	Modesto Ash	BD	FRVE	5	5
Gingko biloba	Gingko	BD	GIBI	19	13
Gleditsia triacanthos	Honey Locust	BD	GLTR	4	4
Koelreutaria paniculata	Goldenrain Tree	BD	KOPA	8	6
Lagerstroemia indica	Crape Myrtle	BD	LAIN	5	0
Liquidambar stryaciflua	Sweetgum	BD	LIST	5	4
Southern Magnolia	Magnolia Grandiflora	BE	MASP	5	5
Pistacia chinensis	Chinese Pistachio	BD	PICH	3	2
Pinus thunbergii	Japanese Black Pine	C	PITH	9	8
Plantanus × acerifolia	London Plane	BD	PLAC	9	9
Prunus cerasifera	Flowering Plum	BD	PRCE	9	6
Pyrus calleryana 'Bradford'	Callery Pear	BD	PYCA	10	9
Quercus ilex	Holly Oak	BE	QUIL	6	6
Zelkova serrata	Zelkova	BD	ZESE	5	4
Total				149	118

BD, broadleaf deciduous; BE, broadleaf evergreen; C, conifer.

field-sampled tree database was created containing 640 trees for these 22 species (18 broadleaf deciduous and 4 evergreen species). All field data were added to the tree GIS database.

High-resolution AVIRIS data were acquired at 11:40 PST (Pacific Standard Time) on 10 October 1998. This low altitude AVIRIS imagery was obtained by a National Oceanic and Atmospheric Administration (NOAA) Twin Otter aircraft flying at an altitude of 3810 m above sea level. At this altitude, AVIRIS pixels have a spatial resolution of about 3.5 m. The high-resolution flight navigation and engineering data (such as GPS and inertial data) recorded during the AVIRIS overflight were used during post-processing to correct for aircraft motion by the Jet Propulsion Laboratory, NASA (Boardman 1998). This AVIRIS dataset covered the central core of the city. Three hundred and forty of the 648 field measured trees in summer 1998 were covered within the AVIRIS flight pass.

Tree species, size and tree type distributions are shown in figure 2. Broadleaf deciduous (BD) trees represent 81% of the sample population in the AVIRIS flight paths. Broadleaf evergreen (BE) and conifer (C) trees represent 14% and 5%, respectively. Due to the loss of on-board global positioning system (GPS) data during the AVIRIS data collection for scenes 5 and 6, our final analysis was focused on the first four scenes of data, thus only 149 field-sampled trees were covered in these four scenes. Table 1 lists tree sample information (species botanical and common names, species code, total and usable samples (crown diameter greater than one pixel) for this study. Tree samples were eliminated if the crown diameter was less than the pixel size.

3.3. *AVIRIS data and spectral analysis*

Because of the relatively small study area and absence of topography, we did not retrieve surface reflectance by correcting the atmospheric scattering and

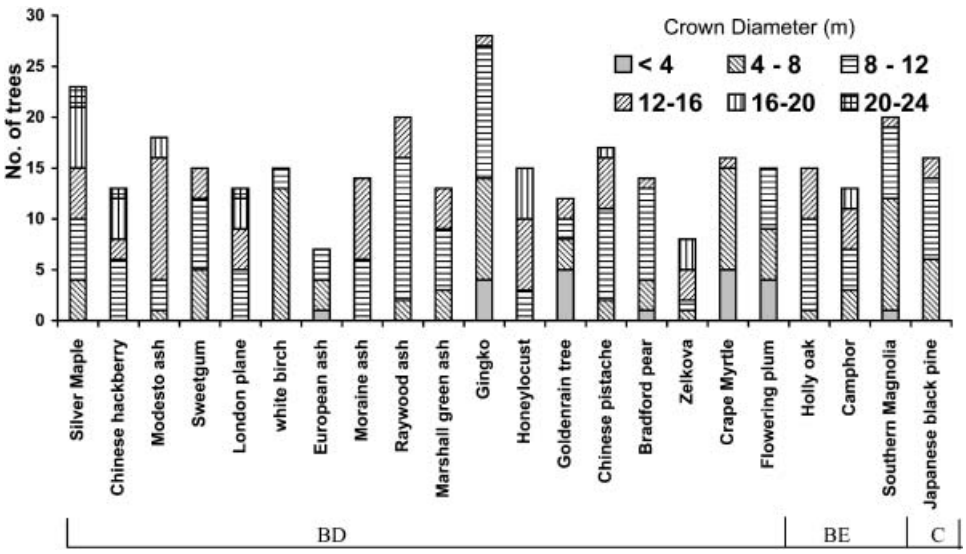


Figure 2. Species distribution of field measured trees. Trees are subdivided into different crown diameter ranges (m) as shown in the figure legend. The tree type is indicated by BD (broadleaf deciduous), BE (broadleaf evergreen) and C (conifer).

absorption. Instead, we assumed that all spectral differences between pixels were due to the surface properties. Radiance data were used for selecting endmembers and for the analyses. Three bands of AVIRIS data (550 nm, 650 nm and 850 nm) were geo-registered to and overlaid with the street layer in the GIS. By overlaying the tree layer in the GIS database with this AVIRIS data, the location of different tree species were identified in the AVIRIS data. Not all field measured trees were selected for extracting spectral information to create the spectral library. We identified a set of criteria that must be satisfied before we used the information. For the site criteria, we focused on the tree crown dimensions and growing environment. The tree crown diameter must be wide enough to cover at least one pixel. We used only isolated trees to avoid possible spectral mixing from other species. Twenty-two species from the field-sampled GIS database were selected. Using their locations we were able to extract spectral reflectance for different tree species and for different surface cover properties. Environment for Visualizing Images, Research Systems Inc., Lafayette, CO, 1997 (ENVI) was used to create the spectral library, masks and to perform the spectral analysis. After we obtained the spectral library for tree species and other ground surface cover properties, we used linear spectral mixture analysis (SMA) to create the layered information for each tree type and species. SMA is based on the assumption that remotely sensed spectral reflectance is a linear summation of the different spectral components and that a specific location or a single pixel can be presented as fractions of endmember types. The endmember fractions of each pixel were used for both tree type and species mapping.

Because mapping accuracy relies on SMA, the selection of endmember for each tree species is important. For each tree species, we identified a set of pixels that expressed the variation in spectral information for that species. We selected the endmember spectra from this set based on the spectrum that was closest to the mean value for each species. We ceased using the set of spectra that expressed the 'within-species' spectral characteristics. The mean and the offset from the mean were calculated for each wavelength. The root mean square (RMS) was calculated for each tree sample. The tree with the smallest RMS value was selected as the endmember for each tree species. For example, there were a total of 19 Gingko trees in the study area for which we had direct field measurements, but six had crown diameters less than one pixel. They were eliminated as potential endmembers. The remaining 13 Gingko (*Gingko biloba*) tree spectra were extracted from the AVIRIS data. The standard deviation from the mean was 171 and the RMS ranged from 73 to 257. The tree with the RMS of 73 was selected as the endmember for the Gingko species.

3.4. *Spectral reflectance characteristic of trees*

Figure 3 shows spectral radiance for the 22 tree species in the spectral library. An offset of 1 μWcm^{-2} nm^{-1} sr^{-1} is used among species for clarity. The spectral difference among species is clearly illustrated in the figure. More detailed spectral radiance is presented in figure 4.

Figure 4(*a*) shows the spectral radiance for three different tree species that represent three different physiognomic types from this AVIRIS data. Japanese black pine (*Pinus thunbergii*), a conifer, has lowest reflected radiance at all wavelengths due to the small intercellular air space inside the leaf and its needle shape. Raywood ash (*Fraxinus oxycarpa 'Raywood'*), a broadleaf deciduous species

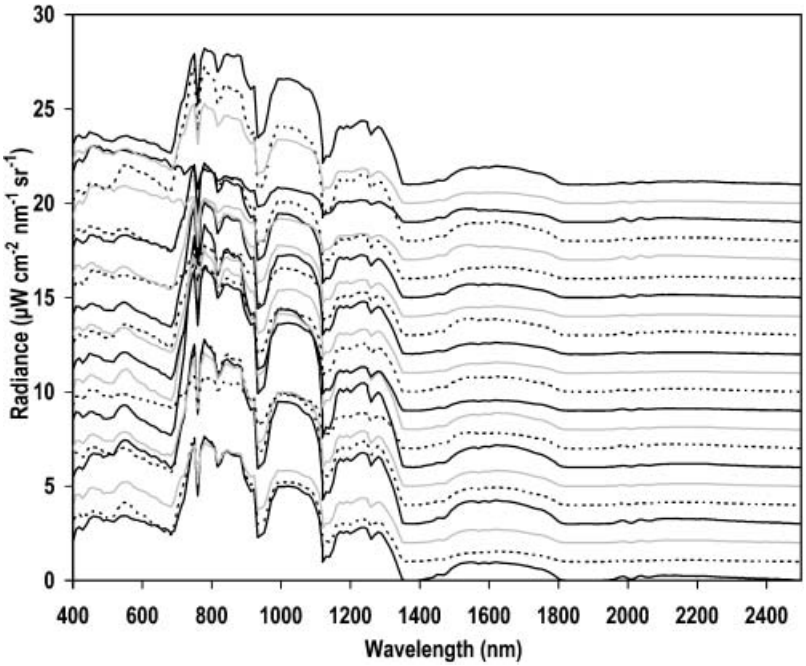


Figure 3. Spectral radiance (μW cm^{-2} nm^{-1} sr^{-1}) of 22 tree species. The tree species from bottom to top are ACSA, BEPE, CESI, CICA, FREX, FROM, FROX, FRPE, FRVG, GIBI, GLTR, KOPA, LAIN, LIST, MASP, PICH, PITH, PLAC, PRCE, PYCA, QUIL, and ZESE. The 1 μW cm^{-2} nm^{-1} sr^{-1} was used as offset for clarity (e.g. for each wavelength, 1 was added to BEPE, 2 was added to CESI, … , and 21 was added to ZESE).

has the highest reflected radiance, and Camphor (*Cinnamomum camphora*), a broadleaf evergreen species is in the middle range. The spectral library, showed as expected, that broadleaf deciduous tree species have higher spectral radiance than broadleaf evergreen trees. Figure 4(*b*) shows the spectral radiance of Modesto ash (*Fraxinus velutina 'Modesto'*), Bradford pear (*Pyrus calleryana 'Bradford'*), Gingko (*Ginkgo biloba*), Zelkova (*Zelkova serrata*), Moraine ash (*Fraxinus holotricha 'Moraine'*), and Goldenrain (*Koelreuteria paniculata*) trees. We only show six species but a similar pattern was observed for the other 12 deciduous species as shown in figure 3. Spectral radiance of these species not only varies in magnitude but also in shape. Among these species, Modesto ash and Bradford pear have higher radiance in the infrared wavelength region. Modesto ash has the highest radiance in the near-infrared region. But in the visible region, it has less radiance than Gingko and Bradford pear; Bradford pear has the highest radiance in the visible region; Gingko and Zelkova have next higher radiance in the near infrared (780 nm) and middle infrared (1000 nm) range. The radiance of Moraine ash trees is higher than Gingko at 950–1100 nm range but reversed at the 800–900nm range. At most wavelengths, the Goldenrain tree has lowest radiance, but at 1300 nm region its radiance was higher than the Gingko.

A common problem faced with this type of data is treating the mixed pixels that are common in an analysis of isolated trees. SMA was used to identify the ground surface properties related to certain land cover types. In practice, the problem of

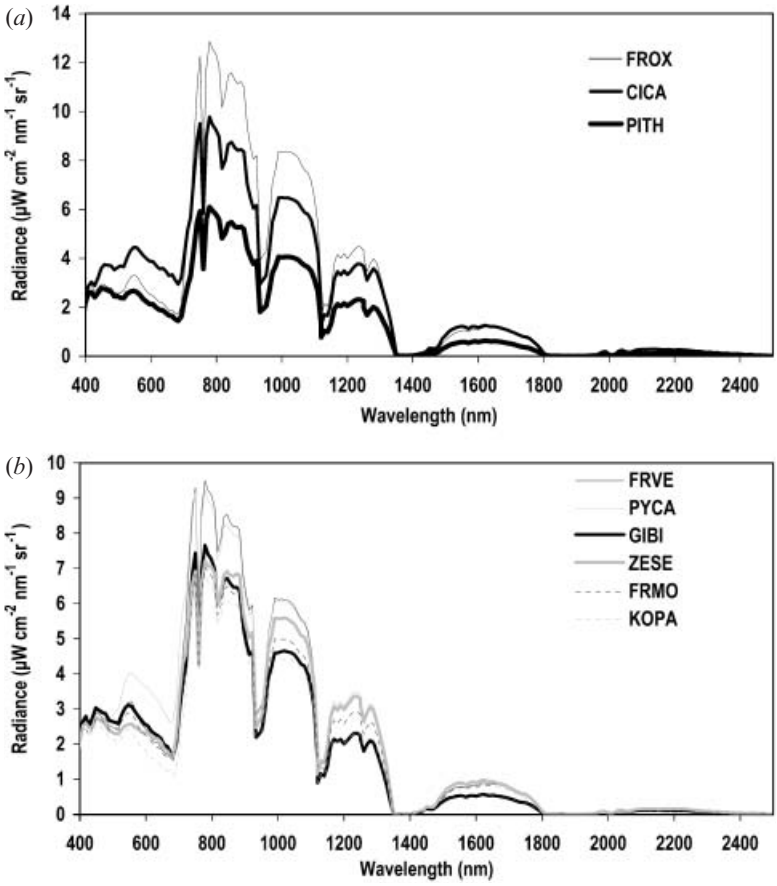


Figure 4. (*a*) Radiance of three tree types. Japanese black pine (conifer) has the lowest radiance. Raywood Ash (broadleaf deciduous) has the highest radiance in the infrared region. Camphor (broadleaf evergreen) has a radiance between the other two trees, but it has the highest radiance in visible wavelength region. (*b*) Radiance of different broadleaf deciduous tree species. Modesto ash and Bradford pear have higher radiance in the infrared wavelength region. But Bradford pear has the highest radiance in the visible region. Goldenrain has the lowest radiance.

uncertainty or noise increases as the number of endmembers increases. In this study, we focused on street trees, thus we restricted the possible combinations of mixing that exist. It is possible to reduce the number of endmembers by masking out areas of non-interest. A multiple-masking technique was used for identifying tree species and for mapping. There are 224 bands in the AVIRIS dataset, but we used only 131 bands. The remaining bands were identified as 'bad bands' (e.g. water absorbing bands and bands with wavelengths greater than 1800 nm). This is due to their low spectral information content for vegetation.

High correlation exists among endmembers of tree species due to the similarity of plant biochemistry. However, the correlation analysis (using the method of least squares) of the endmembers shows that the slope and intercept of the fit are significantly different among each pair of tree species. Thus, SMA method can be used to map and identify the spectral differences of these tree species.

3.5. *Analysis procedures*

Geo-referencing the AVIRIS data was done by using the street layer in the GIS database as reference (Arc/Info, Environmental Systems Research Institute Inc., Redlands, CA, 1997) to locate coordinates and to allow overlaying of other GIS data layers in the database over the low altitude AVIRIS data. Using the street tree layer at GIS database to locate the trees' location in AVIRIS data, a spectral library was manually constructed for the field sampled 22 tree species and for different ground surface cover properties (such as pavement types, buildings and other land cover types). A set of masks was generated based on the SMA analysis and NDVI. The first mask was created based on NDVI value (threshold value = 0.48), which was used for identifying vegetation and non-vegetation as shown in the analysis scheme (figure 5). The flow chart illustrates the sequence of analyses to obtain land cover type, tree type and species identifications. Three types of vegetation were defined (e.g. tree, shrub and turf grass). Masks for these vegetation types were further generated after non-vegetation was masked out. These masks were created based on SMA and the endmember spectra were from the spectral library. Because the SMA analysis yielded a fraction of endmembers in a pixel, the type of tree species was determined by the majority fraction. The same processes were followed until tree type and species were identified. Based on the spectral library developed for the study site that included selected tree species and other land use, different spectral analyses were performed to obtain other data layers that characterized land cover types and tree types (i.e. broadleaf deciduous, broadleaf evergreen and conifer) that contributed to characterizing different tree species. The results were evaluated for accuracy by independent ground reference surveys and by tree identification using the City of Modesto GIS database.

3.6. *Accuracy check*

One of the challenging tasks in remote sensing data processing is to assess accuracy. We chose to use the standard producers, which uses the confusion matrix as the basis for comparison. The accuracy of the classification models was assessed on a tree basis. An error matrix was computed (Card 1982, Davis and Goetz 1990, Congalton 1991), which takes into account the omission and commission errors. We first evaluated the accuracy of all tree species covered by the AVIRIS data. Trees in the GIS database were used as reference. A 3×3 ($10.5\,\text{m} \times 10.5\,\text{m}$) average

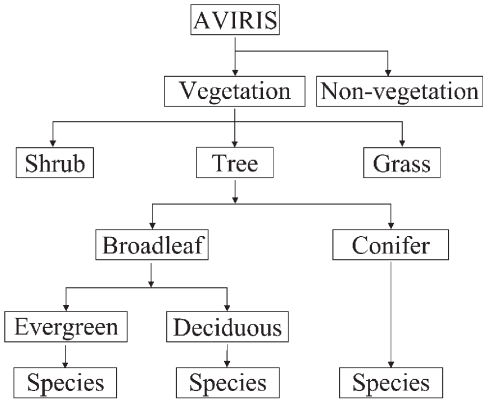


Figure 5. Tree type and species identification flow chart.

window was used to perform the accuracy check because the tree location in the GIS layer was geo-referenced by street parcel address. Use of this window during the accuracy evaluation would reduce the error induced by the initial tree GIS mapping, which was based on street address. In addition to this point assessment, an area assessment for evaluating mapping accuracy was used. We compared all of the street trees in a large city block (2700 m × 2700 m). The results are presented in an error matrix. The overall mapping accuracy was calculated from the number of correctly mapped trees to the total number of trees.

4. Results

We successfully separated land surface cover into different types, such as bare soil, pavement, building, water (not shown) and vegetation by using SMA and multiple masks. The vegetation class was further divided into grass, shrub and tree. The tree class was further identified by physiognomic type (e.g. broadleaf deciduous, broadleaf evergreen and conifer) and by species. Spectral radiances of these endmembers were used in SMA analysis. Because the area had little topographic change (range of 11 m) and the area was so small, an assumption of atmospheric homogeneity over the site was reasonable; we analysed the data in radiance rather than reflectance. This reduced any errors or artefacts due to miscalibration. In this study, we focused on tree characterization so the results are presented in terms of tree types and tree species.

4.1. *Tree type*

Figure 6 shows the original four scenes AVIRIS false colour image (R = 850 nm, G = 650 nm, B = 550 nm) and tree type map. Both the tree and land use patterns are clearly seen on this image (figure 6(*a*)). The conifer tree layer from SMA (figure 6(*b*)) shows where these types of tree were located. The same type of information was obtained for the broadleaf evergreen (figure 6(*c*)) and for broadleaf deciduous trees

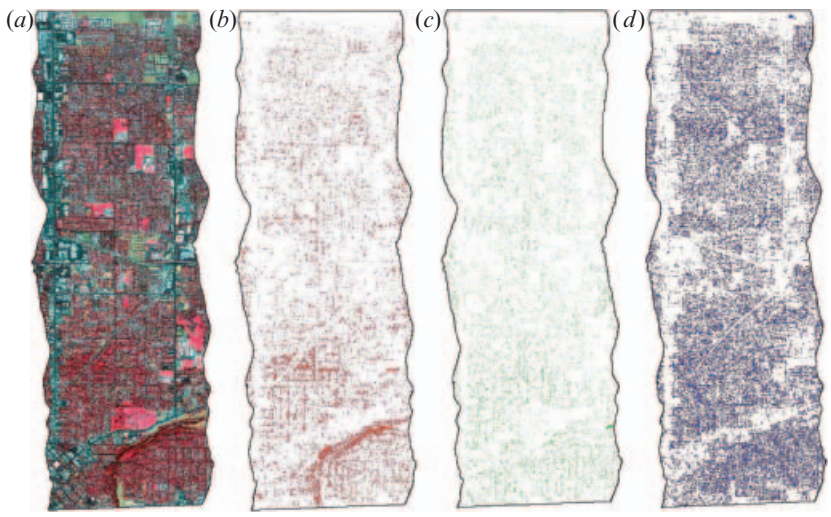


Figure 6. Tree type spatial distribution of study area. (*a*) Colour-infrared Airborne Visible Infrared Imaging Spectrometer (AVIRIS) image (R = 850 nm, G = 650 nm, B = 550 nm). (*b*) Conifer classified pixels in red, (*c*) broadleaf evergreen tree pixels in green, and (*d*) broadleaf deciduous tree pixels in blue.

Table 2. Tree physiognomic type classification error matrix (percentage of total samples matched) of the study site. The rows show the distribution of each type from AVIRIS and the columns show the distribution of the AVIRIS class in the GIS base map. The overall accuracy for all types is 94%.

		GIS base map					
		Conifer		Broadleaf evergreen		Broadleaf deciduous	
Tree type	Total trees	Classified	Trees (%)	Classified	Trees (%)	Classified	Trees (%)
C	8	6	75	0	0	2	25
BE	18	1	6	15	83	2	11
BD	81	1	1	0	0	80	99
Total	107	8	7	15	14	84	79

C, conifer; BE, broadleaf evergreen; BD, broadleaf deciduous.

(figure 6(*d*)). The tree identification accuracy average was 94% (table 2). Average accuracy was 99% for broadleaf deciduous, 83% for broadleaf evergreen and 75% for conifer. Most misidentified trees had small crown sizes with few branches suggesting that subpixel identifications were problematic. In some cases, the average crown diameter was large but irregular. For example, one Japanese black pine tree had a crown diameter 12.6 m in one direction but only 2.4 m in the perpendicular direction. Thus in AVIRIS pixels with this crown, much of the radiance is contributed by the background cover rather than the tree itself. In addition to small tree sizes, some young trees are partially planted under other larger trees. For example, many flowering plum trees are located under larger deciduous trees, based on our field observations. Man-made structures (such as parks, buildings, and streets etc.) can be clearly seen in these figures because the dimensions of these structures are larger than pixels.

4.2. *Tree species*

At the tree species level, our goal was to identify 22 tree species, but here we only examined accuracy for 16 of the 22 species because only these species had statistically valid field samples with tree crown sizes greater than one pixel (table 3). Accuracies were higher for tree species that had relatively large crown sizes and dense leaves, such as the Holly oak, Zelkova and Gingko. Both Modesto ash and Flowering plum had relatively low identification accuracy (20% and 17%). The Modesto ash trees had large crown sizes but fewer branches, resulting in low crown densities which made the spectra of the tree covered pixels fully mixed with land cover underneath the tree crown. Flowering plum trees generally had a relatively small crown diameter and tree height. In addition to these two disadvantages, many Flowering plum trees were planted under other larger tree crowns. Thus, the spectral character of this tree species was strongly affected by adjacent trees when it was measured from the airbone sensor. The spatial distribution of tree species for a portion of the study area is shown in figure 7. In this small area, tree species composition is relatively simple. Zelkova, Gingko, and Hackberry were the dominant street trees. These three tree species accounted for 83% of the total street trees in this area. The rest of the street trees were composed of conifer (5%), broadleaf evergreen (5%) (Magnolia, Holly oak and Camphor), and other broadleaf deciduous trees (7%). A colour infrared composite from AVIRIS data is shown in

Table 3. Tree species classification error matrix shows errors for 16 tree species. The rows show the distribution of each type from AVIRIS and the columns show the distribution of the AVIRIS class in the GIS base map. The overall accuracy for these 16 species is 70%. Species code is found in table 1.

		Species																
Species	Number of trees	ACSA	CICA	FRMO	FRPE	FRVE	GIBI	GLTR	KOPA	LIST	MASP	PITH	PLAC	PRCE	PYCA	QUIL	ZESE	Total
ACSA	6	5											1					6
CICA	7		3	1	1						1	1						7
FRMO	5			4							1							5
FRPE	10				6						1	2				1		10
FRVE	5	2				1						2						5
GIBI	13						13											13
GLTR	4		1					3										4
KOPA	6	2							2			2						6
LIST	4									4								4
MASP	5									1	3	1						5
PITH	8						1			1		6						8
PLAC	9	1							1				6			1		9
PRCE	6						1		1		1	2		1				6
PYCA	9	1						1				2			4	1		9
QUIL	6															6		6
ZESE	4																4	4
Total	107	11	4	5	7	1	15	4	4	6	7	18	7	1	4	9	4	107

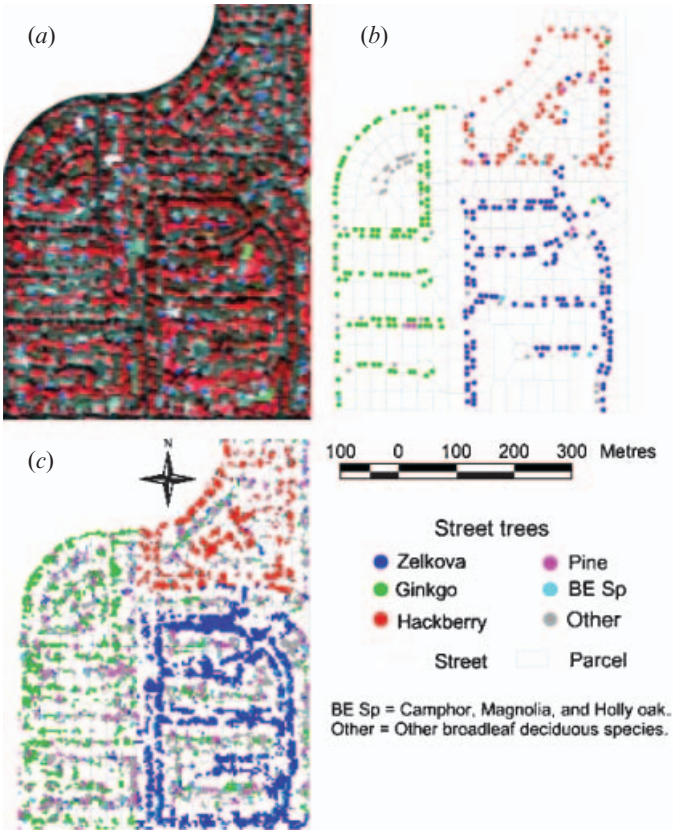
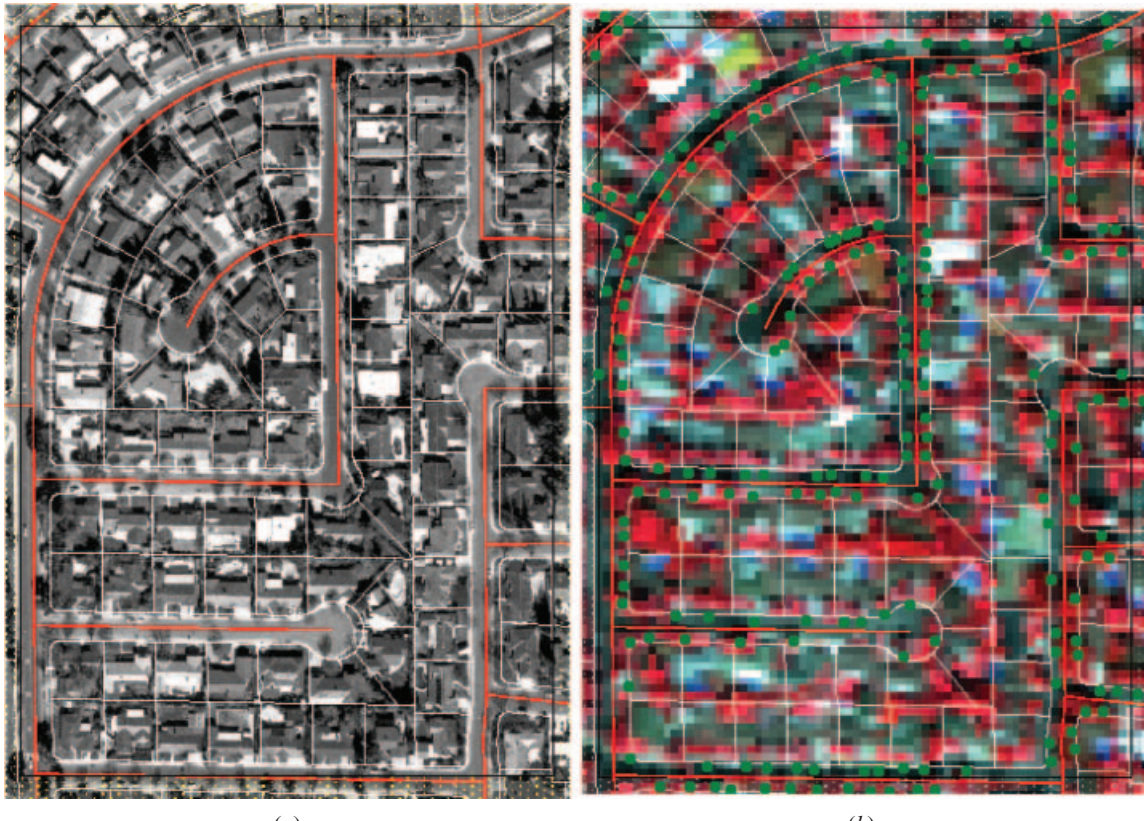


Figure 7. Tree species spatial distribution of a selected area of Modesto. (*a*) Colour infrared AVIRIS image, (*b*) tree species identification from the GIS database, and (*c*) tree species identification derived from AVIRIS data.

figure 7(*a*). The spatial distribution of tree species in the GIS layer is shown in figure 7(*b*). The result of the classification of tree species from the AVIRIS analysis is presented in figure 7(*c*). It is clear that figure 7(*b*) and (*c*) are in close agreement, except that figure 7(*c*) includes tree information for non-street trees and includes more precise information on spatial distribution. Figure 7(*b*) only includes the street trees that were maintained by the city.

The AVIRIS data at 3.5 m pixels had a rather high spectral resolution, which provided rich and unique spectral information for each tree species. However, the spectral mixing that occurred from overlapping foliage and shadows from large trees on smaller ones restricted the accuracy of tree identification. In contrast to trees in a rural forest, urban forest trees are sparse and often open-grown. Species can vary widely from tree to tree, thus high spatial resolution is also import for urban tree species mapping. A high spatial resolution (0.5 m spatial resolution) grey-scale aerial photograph and a colour-infrared image from three AVIRIS bands are shown in figure 8. It is easy to identify tree locations and separate individual trees in the aerial photograph. However, the limited spectral information of this grey-scale photograph is not easily translated into a species map. The hyperspectral AVIRIS dataset has a rich depth of spectral information but most tree pixels are mixed at the



(*a*) (*b*)

Figure 8. Tree location on aerial photograph (*a*) and AVIRIS image (*b*). The high spatial resolution aerial photograph (0.5 m) allows for easy interpretation of tree locations from address referencing.

3.5 m pixel resolution. Increasing the spatial resolution of hyperspectral data would improve the accuracy of urban tree mapping.

5. Conclusions and discussion

Our analyses demonstrate that isolated tree species can be identified and separated with high accuracy by type using high spatial resolution AVIRIS data. We used a GIS database to identify training sites and to validate the final maps. In addition to tree characterization, these data can be used for characterizing land cover. For example, we can separate the man-made structures by the materials that are used, such as different types of buildings, houses, concrete pavement and asphalt pavement. The potential value of these data for urban forest applications includes estimating tree health (e.g. evidence for stress) and leaf area for different tree species and site conditions. AVIRIS data acquired in spring or summer rather than October might provide better separation of some species or additional information about tree condition. For example, data acquired in both summer and winter seasons could be used to easily identify locations of deciduous and evergreen trees.

The multiple-masking techniques used in this study show an improvement in tree identification accuracy compared to identification without using this technique (Xiao *et al*. 1999). The mixing of land cover for street trees is relatively simple to characterize. For example, pixels of most street trees in residential areas are mixed with road and/or turf grass. Street trees were also mixed with bare soil and/or road in median strips and in some commercial areas. This mixing reduces the number of possible endmembers. This is the greatest reason that accuracy increased compared to our earlier results. Using this method to identify trees in locations other than along the street may not yield the same results due to the potential for more complex mixing combinations.

The AVIRIS data had high spectral resolution (224 bands) and relatively high spatial resolution ($\sim$3.5 m), which provided a rich and unique set of spectral information for each tree species. However, spatial resolution is at least as important as spectral resolution for urban tree species mapping, because most trees were still within mixed pixels at this scale. Increasing spatial resolution of the hyperspectral dataset could improve the accuracy of tree identification.

Acknowledgments

The authors would like to thank the Engineering and Transportation Department of the City of Modesto, California for providing the basic GIS data layers. We appreciate Paula Peper at the Center for Urban Forest Research, USDA Forest Service, Pacific Southwest Research Station, Davis, CA, and Steven P. Lennartz at University of New Hampshire, Durham, NH, for their assistance in field data collection and data analysis. This research was supported in part by funds provided by the Pacific Southwest Research Station, USDA Forest Service.

References

AMBROSIA, V. G., BUECHEL, S. W., BRASS, J. A., and PETERSON, J. R., 1998, An integration of remote sensing, GIS, and information distribution for wildfire detection and management. *Photogrammetric Engineering and Remote Sensing*, **64**, 977–985.

ARNOLD, C. L., and GIBBONS, C. J., 1996, Impervious surface coverage—the emergence of a key environmental indicator. *Journal of the American Planning Association*, **62**, 243–258.

ARTHUR-HARTRANFT, S. T., CARLSON, N. T., and CLARKE, K. C., 2003, Satellite and ground-based microclimate and hydrologic analyses coupled with a regional urban growth model. *Remote Sensing of Environment*, **86**, 385–400.

AVERY, T. E., and BERLIN, G. L., 1992, *Fundamentals of Remote Sensing and Airphoto Interpretation*, 5th edn (Upper Saddle River, NJ: Prentice-Hall).

BLACKBURN, G. A., and MILTON, E. J., 1997, An ecological survey of deciduous woodlands using airborne remote sensing and geographical information system (GIS). *International Journal of Remote Sensing*, **18**, 1919–1935.

BOARDMAN, J., 1998, Precision geocoding of low altitude AVIRIS data: lessons learned in 1998. *Summaries of the Eighth JPL Airborne Earth Science Workshop*, *Pasadena CA* (Pasadena, CA: JPL), pp. 63–68.

CARD, D. H., 1982, Using known map category marginal frequencies to improve estimates of thematic map accuracy. *Photogrammetric Engineering and Remote Sensing*, **48**, 431–439.

CLAPHAM W. B., 2003, Continuum-based classification of remotely sensed imagery to describe urban sprawl on a watershed scale. *Remote Sensing of Environment*, **86**, 322–340.

CONGALTON, R., 1991, Review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, **37**, 35–46.

DAVIS, F. W., and GOETZ, S., 1990, Modeling vegetation pattern using digital terrain data. *Landscape Ecology*, **4**, 69–80.

DOUGLAS, I., 1983, *The Urban Environment* (London: Edward Arnold).

DWYER, J. F., MCPHERSON, E. G., SCHROEDER, H. W., and ROWNTREE, R. A., 1992, Assessing the benefits and costs of the urban forest. *Journal of Arboriculture*, **18**, 227–234.

EHLERS, M., 1990, Remote sensing and geographic information system: towards integrated spatial information processing. *IEEE Transactions on Geoscience and Remote Sensing*, **28**, 763–766.

FRANKLIN, S. E., WULDER, M. A., and GERYLO, G. R., 2001, Texture analysis of IKONOS panchromatic data for Douglas-fir forest age class separability in British Columbia. *International Journal of Remote Sensing*, **22**, 2627–2632.

FRANKLIN, S. E., HALL, R. J., SMITH, L., and GERYLO, G. R., 2003, Discrimination of conifer height, age and crown closure classes using Landsat-5 TM imagery in the Canadian Northwest Territories. *International Journal of Remote Sensing*, **24**, 1823–1834.

GALLO, K. P., and TARPLEY, J. D., 1996, The comparison of vegetation index and surface temperature composites for urban heat-island analysis. *International Journal of Remote Sensing*, **17**, 3071–3076.

GONG, P., and HOWARTH, P. J., 1990, The use of structure information for improving land-cover classification accuracies at the rural–urban fringe. *Photogrammetric Engineering and Remote Sensing*, **56**, 67–73.

GREEN, R. O., EASTWOOD, M. L., SARTURE, C. M., CHRLEN, T. G., ARONSSON, M., CHIPPENDALE, B. R., FAUST, J. A., PAVRI, B. E., CHOVIT, C. J., and SOLIS, M., 1998, Imaging spectroscopy and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). *Remote Sensing of Environment*, **65**, 227–248.

GRIMMOND, C. S. B., and SOUCH, C., 1994, Surface description for urban climate studies: a GIS based methodology. *Geocarto International*, **9**, 47–59.

HUANG, M. C., and RIDD, M. K., 2002, A subpixel classifier for urban land-cover mapping based on a maximum-likelihood approach and expert system rules. *Photogrammetric Engineering and Remote Sensing*, **68**, 1173–1180.

HUANG, C., and TOWNSHEND, J. R. G., 2003, A stepwise regression tree for nonlinear approximation: applications to estimating subpixel land cover. *International Journal of Remote Sensing*, **24**, 75–90.

HUANG, X. M., LYONS, T. J., SMITH, R. C. G., and HACKER, J. M., 1995, Estimation of land surface parameters using satellite data. *Hydrological Processes*, **9**, 631–643.

LAKSHMI, V. U., MURTHY, M. S. R., and DUTT, C. B. S., 1998, Efficient forest resources management through GIS and remote sensing. *Current Science*, **75**, 272–282.

LI, X., 1998, Measurement of rapid agricultural land loss in the Pearl River Delta with the integration of remote sensing and GIS. *Environment and Planning B-Planning and Design*, **25**, 447–461.

McBride, J. R., and Jacobs, D. F., 1986, Presettlement forest structure as a factor in urban forest development. *Urban Ecology*, **9**, 245–266.

McPherson, E. G., Nowak, D., Heisler, G., Grimmond, S., Souch, C., Grant, R., and Rowantree, R., 1997, Quantifying urban forest structure, function, and value: the Chicago Urban Forest Climate Project. *Urban Ecosystems*, **1**, 49–61.

Martin, L. R. G., Howarth, P. J., and Holder, G., 1988, Multispectral classification of land use at the rural–urban fringe using SPOT data. *Canadian Journal of Remote Sensing*, **14**, 72–79.

Morgan, K. M., Newland, L. W., Weber, E., and Busbey, A. B., 1993, Using SPOT satellite data to map impervious cover for urban runoff predictions. *Toxicological and Environmental Chemistry*, **40**, 11–16.

Nowak, D. J., Rowntree, R. A., McPherson, E. G., Sisinni, S. M., Kerkmann, E. R., and Stevens, J. C., 1996, Measuring and analyzing urban tree cover. *Landscape and Urban Planning*, **36**, 49–57.

Peper, P. J., and McPherson, E. G., 2003, Evaluation of four methods for estimating leaf area of isolated trees. *Urban Forestry and Urban Greening*, **2**, 19–29.

Peper, P. J., McPherson, E. G., and Mori, S. M., 2001, Equations for predicting diameter, height, crown width, and leaf area of San Joaquin Valley street trees. *Journal of Arboriculture*, **27**, 306–317.

Price, K. P., Guo, X. L., and Stiles, J. M., 2002, Optimal Landsat TM band combinations and vegetation indices for discrimination of six grassland types in eastern Kansas. *International Journal of Remote Sensing*, **23**, 5031–5042.

Rowntree, R. A., 1984, Ecology of the urban forest—introduction to Part 1. *Urban Ecology*, **8**, 1–11.

Shao, G. F., Young, D. R., Porter, J. H., and Hayden, B. P., 1998, An integration of remote sensing and GIS to examine the responses of shrub thicket distributions to shoreline changes on Virginia Barrier Islands. *Journal of Coastal Research*, **14**, 299–307.

Underwood, E., Ustin, S., and DiPietro, D., 2003, Mapping nonnative plants using hyperspectral imagery. *Remote Sensing of Environment*, **86**, 150–161.

USDA, 1975, Urban Hydrology for Small Watersheds. US Department of Agriculture, Technical Release 55, PB82140666.

Ustin, S. L., and Xiao, Q. F., 2001, Mapping successional boreal forests in interior central Alaska. *International Journal of Remote Sensing*, **22**, 1779–1999.

Weber, C., and Puissant, A., 2003, Urbanization pressure and modeling of urban growth: example of the Tunis Metropolitan Area. *Remote Sensing of Environment*, **86**, 341–352.

World Population Reference, 1993, World population situation, United Nations, New York, USA.

World Resources Institute, 1996, *World Resources — The Urban Environment* (New York: Oxford University Press).

Xiao, Q. F., and McPherson, E. G., 2003, Rainfall interception by Santa Monica's municipal urban forest. *Urban Ecosystems*, **6**, 291–302.

Xiao, Q. F., Ustin, S. L., McPherson, E. G., and Peper, P. J., 1999, Characterization of the structure and species composition of urban trees using high resolution AVIRIS data. *Summaries of the Eighth JPL Airborne Earth Science Workshop*, *Pasadena CA* (Pasadena, CA: JPL), pp. 451–460.

Xu, B., Gong, P., and Pu, R. L., 2003, Crown closure estimation of oak savannah in a dry season with Landsat TM imagery: comparison of various indices through correlation analysis. *International Journal of Remote Sensing*, **24**, 1811–1822.